Coherent change detection in repeat-pass synthetic aperture sonar

Blair Bonnett, M.E., B.E. (Hons)

A thesis presented for the degree of
Doctor of Philosophy
in
Electrical and Computer Engineering
at the
University of Canterbury,
Christchurch, New Zealand.

March 2017

ABSTRACT

Change detection involves comparing two images to find the differences between them. With sonar data, this can be used to identify changes to both the seafloor and objects located on it. As well as mine-hunting applications, it has been used to spot changes made to the seafloor by bottom-feeding fish, detect the effects of sediment transport on the surface topography of the seafloor, and identify changes to the water properties.

The ability to perform change detection is dependent on the images from the two passes of the sonar being precisely registered (aligned). The navigation data alone is not sufficiently accurate to perform this registration. Therefore, the registration must be corrected via a data-driven approach, estimating the errors from the data itself. Registering images of a bland seafloor has been described as the worst-case scenario due to the lack of distinguishing features to match between the images.

This thesis presents a number of modifications to an existing rough facet synthetic aperture sonar (SAS) simulator to allow the generation of suitable repeat-pass imagery. The implementation of the scattering model is extended to allow the temporal decorrelation of a facet between passes to be simulated. These repeat-pass images are validated against published statistical models for the coherence of two speckle images.

As with any coherent imaging system, SAS imagery is corrupted by multiplicative speckle noise. Despite its random appearance, speckle is a deterministic process. If the images are aligned, then the speckle patterns in corresponding areas are coherent (assuming the temporal decorrelation is sufficiently low). This implies that the speckle patterns in two images can be used to register them. A model of the position errors caused by incorrect track information is presented, along with a block-based correlation method to estimate these errors from two images. Using these detected errors, parameter estimation can be used to find the translation and heading errors in the assumed sonar track.

Bland repeat-pass images generated by the simulator are used to test the performance of this registration algorithm. It is shown to be robust against both additive noise and temporal decorrelation, and capable of detecting the track errors. Guidelines are provided for both the design of the system and the implementation of the presented algorithm in order to maximise the registration accuracy.

Deputy Vice-Chancellor's Office
Postgraduate Office

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

Co-Authorship Form

This form is to accompany the submission of any thesis that contains research reported in co-authored work that has been published, accepted for publication, or submitted for publication. A copy of this form should be included for each co-authored work that is included in the thesis. Completed forms should be included at the front (after the thesis abstract) of each copy of the thesis submitted for examination and library deposit.

Publication 1: B. Bonnett and M. P. Hayes [2016]. "Simulation of temporal coherence loss for repeat-pass synthetic aperture sonar". In: *2016 International Conference on Image and Vision Computing New Zealand*. IVCNZ '16. Palmerston North, New Zealand.
Parts of sections 4.3 (Simulating decorrelated facets), 5.1 (Goodness of fit testing), 5.2 (Single facet simulations), and 5.3 (Whole scene simulations) present this material.

Publication 2: B. Bonnett and M. P. Hayes [2014]. "Data-driven Image Registration for Coherent Change Detection of Synthetic Aperture Sonar Imagery". In: *Proceedings of the 29th International Conference on Image and Vision Computing New Zealand*. IVCNZ '14. Hamilton, New Zealand: ACM, pp. 196–201.
Parts of sections 6.4 (System Model), 6.5 (Correlation peaks), 6.6 (Estimation of model parameters) and 6.8 (Translation errors) present this material.

Publication 3: B. Bonnett, M. P. Hayes, and A. Hunter [2013]. "Registration of images from a hull mounted, low frequency synthetic aperture sonar". In: *28th International Conference on Image and Vision Computing New Zealand*. IVCNZ 2013. Wellington, New Zealand, pp. 142–147.
Section 6.3 (The MUD sonar) presents this material.

Please detail the nature and extent (%) of contribution by the candidate:

Publication 1: 80 % of the research and analysis and 95 % of the writing was contributed by the candidate.

Publication 2: 80 % of the research and analysis and 95 % of the writing was contributed by the candidate.

Publication 3: 75 % of the research and analysis and 95 % of the writing was contributed by the candidate.

CERTIFICATION BY CO-AUTHORS

If there is more than one co-author then a single co-author can sign on behalf of all. The undersigned certifies that:

- The above statement correctly reflects the nature and extent of the PhD candidate's contribution to this co-authored work
- In cases where the candidate was the lead author of the co-authored work, he or she wrote the text.

Name: *Michael Hayes* Signature: MPH Date: *31 March 2017*

ACKNOWLEDGEMENTS

First and foremost, I need to thank my supervisors Associate Professor Michael Hayes and Emeritus Professor Peter Gough for everything they've done to support me over the years. Peter offered me the project shortly before retiring, and in between skiing, climbing, and road-tripping around North America has popped in to see how things were going. Michael's influence on this research has been immense. His ability to explain difficult concepts, decipher weird results, demonstrate how to break down and methodically solve a problem, and correct my abuse of the English language has been vital.

Undertaking a big project like this isn't always a straightforward path. There are plenty of ups and downs along the way. To Debs, Louise, Lance, Chrissy, Michael and Francine, Laura and all the others who've sat down, listened, and helped me talk through the challenges: thanks. The perspective you've provided and the love and support you've shown is hugely appreciated.

Thanks to everybody who makes the department an awesome place to work. Jac and the technical staff have kept the place running through all the disruptions of having the wing rebuilt, and still had time to chat and help postgrads. A special mention is required to Dave van Leeuwen for keeping the computing cluster running despite my best attempts to crash it with dodgy software. Thanks to the academic staff who've been there to help problem-solve and talk about anything and everything. Peter Smith deserves singling out for his tireless efforts in explaining statistics to me. Also Philippa and the various admin staff — Helen, Olivia, Deborah, Catherine — who have to deal with all the paperwork associated with research (in my case, the paperwork is usually late or non-existent) and still manage to keep smiling.

Thanks to all the postgrads who've shared an office / tearoom / building with me and have participated in all the morning and afternoon teas, Tandoori missions, dinners, cake club, and all the other myriad events that have provided a welcome break from the frustrations. Spending time as the postgrad student rep was a privilege.

I've been fortunate enough to have been encouraged to do all sorts of teaching, tutoring, demonstrations and outreach work during my time as a postgrad. Much of this has been in areas outside of my knowledge and comfort zone, and has really pushed me to keep learning and exploring. A special shout-out to to Paul and the HV lab team who

insisted I spent a sufficient amount of time out in the high voltage lab demonstrating to people the uses of 100 kV for research / cooking marshmallows / blowing things up. The many procrastination projects I mucked around on may have frustrated my supervisors, but provided an important break and also helped expand my knowledge. Lance rates another mention here for all the great times making contraptions ranging from an electric scooter/sidecar combination to a tennis ball cannon. The willingness of the technical staff to lend their time and knowledge to these endeavours is also much appreciated.

The University of Canterbury Club (still the Staff Club to many) has featured prominently in my life over the past few years. The opportunity to socialise with people from all departments and areas of campus life, as well as those out in the real world, has been amazing. Getting involved on the Club committee and with the organisation of the South Island Beer Festivals has been a blast. I was a bit disconcerted that after 39 years as manager, Charles retired on the same night I submitted this thesis. Hopefully it wasn't something I said. All the folks I've met through the Society of Beer Advocates and the Christchurch Home Brewers Association need acknowledgement too. Thanks for all the friendship, good times, contacts in the local brewing scene, and help and encouragement to keep improving my own brewing skills.

I would like to thank the University of Canterbury for the financial support received via a University of Canterbury Doctoral Scholarship. This work was started as part of US Office of Naval Research grant N62909-11-1-7037 (change detection using multiple pass InSAS) and I am grateful for the support this provided, both financially and in terms of collaboration and extra data made available for testing. The real-world data analysed in this thesis was generously made available by the TNO and the Royal Netherlands Navy.

Thanks to Dr Thomas Fickenscher and Dr Daniel Sternlicht for examining this work. I am grateful for the time and effort they put into reviewing it and for the feedback they gave. Rick Millane also deserves a mention here for organising and chairing my defense.

And last but most definitely not least, to my family. I know putting up with a perpetual student wasn't always a walk in the park, but I couldn't have done this without your ever-present love, encouragement, and support.

PREFACE

This research began with the ultimate goal of turning change detection with synthetic aperture sonar (SAS) imagery into a solved problem. As may be expected with any such research problem, it turned out to not be as straightforward as that.

It soon became apparent that there were a variety of change detection algorithms out there, many of which had been applied to sonar images. These have been shown to work well, provided the input data is suitable (the "garbage in, garbage out" scenario). In this case, the main requirement is for the images to be accurately registered or aligned: the general rule of thumb is that the misregistration needs to be below one-tenth of the resolution of the sonar. As the navigation data is generally too inaccurate to generate images meeting this requirement, this needs to be corrected based on the images themselves.

Following on from this is the need to be able to validate any registration procedure; a *ground truth* or known answer is required. With real data this forms a circular problem as, if the correct solution was known, the registration step would be unneeded. This is where simulated data is useful. The simulation can be controlled to provide a given scenario (errors in the path the sonar follows, changes to the imaged scene, etc.) along with the ground truth needed to check the behaviour of any registration algorithm applied to the data.

A SAS simulator based on a statistical rough scattering model had been previously developed in the Electrical and Computer Engineering Department at the University of Canterbury, and was shown to generate valid single-pass data. Therefore, the first step in this research was to check its suitability for repeat-pass imaging. Considerable care is required with both simulation and reconstruction processes to ensure coherence is maintained between passes. It soon became apparent that the implementation of the scattering model was not capable of guaranteeing the same realisation of the scene when imaged from a different track. The cause of this was subtle and took a considerable time to isolate and correct.

Following this it was necessary to simulate a particular level of change in a given area of the imaged scene. This involved the development of an extension to the scattering model. Details on repeat-pass simulation are given in Chapter 4 and validation of the extended scattering model is presented in Chapter 5.

The limiting case for change detection with sonar imagery is a bland seafloor, i.e., one devoid of any distinguishing features such as rocks. Using the images themselves to correct registration errors first requires estimating the displacement between the images. If stable features are present they provide good reference points to use for alignment. In the bland case it is necessary to compare patches of the seafloor itself. The major contribution of this research is a study into the registration of bland seafloors. In Chapter 6 a simple model of the errors in the repeat-pass track is presented, and a correlation-based registration procedure is evaluated for a range of translation and rotation errors.

Many of the procedures involved in this research required repeated evaluations over a sliding window of some metric. Integral images have been used widely in image processing to provide an efficient means of performing this evaluation. In the sonar field, they have been previously used in the calculation of the lacunarity of a scene. A minor contribution of this research is the extension of this for use with coherence calculations and a study of the performance improvements they provide; this is presented in Chapter 3.

An overview of the layout of this thesis and the assumed background knowledge of the reader is available in the introduction.

PUBLICATIONS

To date, the following works have been published regarding the research presented in this thesis:

B. Bonnett, M. P. Hayes, and A. Hunter [2013]. "Registration of images from a hull mounted, low frequency synthetic aperture sonar". In: *28th International Conference on Image and Vision Computing New Zealand*. IVCNZ 2013. Wellington, New Zealand, pp. 142–147

B. Bonnett and M. P. Hayes [2014]. "Data-driven Image Registration for Coherent Change Detection of Synthetic Aperture Sonar Imagery". In: *Proceedings of the 29th International Conference on Image and Vision Computing New Zealand*. IVCNZ '14. Hamilton, New Zealand: ACM, pp. 196–201

B. Bonnett and M. P. Hayes [2016]. "Simulation of temporal coherence loss for repeat-pass synthetic aperture sonar". In: *2016 International Conference on Image and Vision Computing New Zealand*. IVCNZ '16. Palmerston North, New Zealand

The first two cover content presented in Chapter 6, while the latter covers content from Chapters 4 and 5.

NOMENCLATURE

This section defines the mathematical notation, variables, and acronyms used throughout this thesis. Where possible, standard notation as used in the associated fields is employed. However, this is not always possible due to conflicts between different areas of the research. Additionally, there are often multiple notations in use in different publications covering the same material.

MATHEMATICAL NOTATION

A random variable is capitalised, X, and a realisation of this random variable is lower-case, x. The PDF is denoted $f_X(x)$, with any parameters of the distribution being included as $f_X(x; x_0)$. Similarly, the CDF is represented by $F_X(x)$ or $F_X(x; x_0)$. The expected value is $\mathrm{E}[X]$ and the variance is $\mathrm{Var}[X]$.

Matrices are upper-case and typeset in bold, $\mathbf{A}$, while vectors are lower-case and bold, $\mathbf{b}$. The magnitude (Euclidean norm or L^2 norm) is given by $|\mathbf{b}|$ while a normalised (or unit) vector is $\hat{\mathbf{b}} = \mathbf{b}/|\mathbf{b}|$. The transpose is $\mathbf{A}^{\mathrm{T}}$ and the Hermitian transpose or conjugate transpose $\mathbf{A}^{\mathrm{H}}$ is found by taking the transpose and then calculating the complex conjugate of each entry. The Moore-Penrose pseudoinverse is given by $\mathbf{A}^{+} = (\mathbf{A}^{\mathrm{H}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{H}}$. The dot product of two vectors $\mathbf{x}$ and $\mathbf{y}$ is denoted by $\mathbf{x} \cdot \mathbf{y}$ while their cross-product is $\mathbf{x} \times \mathbf{y}$. The direction cosine $\cos(\mathbf{x}, \mathbf{y})$ is the cosine of the angle between two vectors.

The gamma function $\Gamma(z)$ is an extension of the factorial function to real and complex numbers. The generalised hypergeometric function ${}_pF_q(a_1, \ldots, a_p; b_1, \ldots, q; z)$ is an analytic function based on power series.

MATHEMATICAL VARIABLES

The following sections list major variables used in the different areas of work covered by this thesis. Minor variables — such as those used in a definition — are not given here, and are often reused. Where there is some reuse (for example, D could be either the size of a transducer aperture, or the random variable representing the distribution of the estimated degree of coherence), it is hoped that the context will clarify which

version is in use. Many of these variables may appear with subscripts identifying a particular instance, for example, $B_h(\hat{\mathbf{x}}, f)$ represents the beampattern of the hydrophone or receiver.

Sonar properties

B	Bandwidth of the transmitted pulse
τ_p	Duration of the transmitted pulse
δ_r	Resolution of the system in the across-track (range) direction
δ_y	Resolution of the system in the along-track direction
D	Along-track size of the transducer aperture

Wave propagation and simulation

f	Frequency
c	Speed of propagation; $c \approx 1500\,\mathrm{m/s}$ for acoustic waves in salt water
k	Angular wavenumber, $k = 2\pi f/c$
$a(\mathbf{x})$	Aperture function describing the shape of a transducer or facet
$h(\mathbf{x})$	Height function; describes the height of a rough facet
$B(\hat{\mathbf{x}}, f)$	Beampattern of the transmitted energy
$S(f)$	Signal spectrum
ψ	Acoustic field radiated by the sonar and scattered by the scene
r	Range

Coherence

Λ	Ensemble coherence
δ	Degree of coherence, $\delta = \lvert\Lambda\rvert$
ϕ_0	Coherence phase difference, $\phi_0 = \arg\Lambda$
γ	Spatial coherence estimate
d	Estimated degree of coherence, $d = \lvert\gamma\rvert$
D	Random variable from which d is drawn
ϕ	Estimated phase difference, $\phi = \arg\gamma$
Φ	Random variable from which ϕ is drawn
$\Delta\phi$	Phase correction term applied to the spatial coherence estimate
γ_n	Coherence factor (loss of coherence due to) additive acoustic noise
γ_s	Coherence factor from footprint shift and image misregistration
γ_b	Coherence factor due to baseline decorrelation

γ_p	Coherence factor from processing noise (including grating lobes)
γ_t	Coherence factor due to temporal decorrelation

Registration

M	Number of passes to be registered
P_m	Number of pings in the m^{th} pass
$\mathbf{c}_m$	Centre of the linear approximation to the m^{th} track
$\hat{\mathbf{c}}_m$	Estimated value of $\mathbf{c}_m$
$\mathbf{d}_m$	Displacement vector of the linear approximation to the m^{th} track
$\hat{\mathbf{d}}_m$	Estimated value of $\mathbf{d}_m$
$\mathbf{c}$	Origin of the common reconstruction grid
$\mathbf{h}$	Heading of the y-axis of the common reconstruction grid
$\mathbf{x}_m(p)$	Position of the sonar at the p^{th} ping of the m^{th} pass
$\hat{\mathbf{x}}_m(p)$	Position of the sonar at the p^{th} ping of the linear approximation to the m^{th} track
$\mathbf{x}'_m(p)$	The position $\mathbf{x}_m(p)$ transformed to the common reconstruction grid
t_x	Across-track translation error in the navigation data
t_y	Along-track translation error in the navigation data
α	Heading error in the navigation data
$\mathbf{e}$	Error vector $[t_x, t_y, \alpha]^{\text{T}}$
$\widehat{\mathbf{e}}$	Estimated value of $\mathbf{e}$
Δx	Across-track displacement of a point in the image due to misregistration
Δy	Along-track displacement of a point in the image due to misregistration
$\Delta\mathbf{x}$	Displacement vector $[\Delta x, \Delta y]^{\text{T}}$
$\Delta\widetilde{x}_n$	Observed across-track displacement in the n^{th} block of an image pair
$\Delta\widetilde{y}_n$	Observed along-track displacement in the n^{th} block of an image pair
β	Rotation angle between corresponding blocks from two passes
$\hat{\beta}$	Rotation angle that maximises the correlation between corresponding blocks from two passes

ABBREVIATIONS

AASR	Along-track ambiguity to signal ratio
AUV	Autonomous underwater vehicle
AWGN	Additive white Gaussian noise

CCA	Canonical correlation analysis
CDF	Cumulative distribution function
CRLB	Cramér-Rao lower bound
DVL	Doppler velocity log
EDF	Empirical distribution function
ENL	Effective number of looks
GLONASS	Global navigation satellite system
GPS	Global positioning system
IMU	Inertial measurement unit
LBL	Long baseline (network)
PDF	Probability density function
RMS	Root mean square
RTK	Real time kinematic (satellite navigation)
SAR	Synthetic aperture radar
SAS	Synthetic aperture sonar
SIFT	Scale-invariant feature transform
SNR	Signal to noise ratio
SURF	Speeded up robust features

CONTENTS

Chapter 1

INTRODUCTION

With more than 70 % of the earth's surface covered by water, the ability to map and monitor the underwater environment is desirable scientifically, commercially, and militarily. Sonar imaging is generally the only feasible technique of imaging the seafloor. In most conditions, optical images are difficult to obtain due to the attenuation of electromagnetic energy by the high conductivity of seawater. In many areas of interest the water will also contain suspended sediment, further hindering photographic methods. Sonar systems can be designed to image objects buried beneath the seafloor [Jonsson, Pihl, and Aklint 2005; Nakamura, Yamaguchi, and Hama 2004; Tinkle and Chang 2001] allowing the detection and monitoring of sub-surface objects.

Change detection, the practice of comparing two images and identifying regions that differ, has been used with radar imagery for decades. In recent years there has been interest in applying it to sonar images, and a number of results have been published doing so. With approximately 90 % of world trade transported by sea [Coiras et al. 2008], the detection of mines or other explosive devices in harbours and shipping lanes has been the main application of sonar change detection to date. These locations tend to have large amounts of debris that can resemble potential threats. Applying change detection to regular surveys of the area highlights new objects and allows previously investigated objects to be ignored [Sternlicht, Harbaugh, and Nelson 2009]. Other studies have found changes made by bottom-feeding fish and other seafloor fauna. A number of environmental processes can modify the seafloor, for example, sediment transport by wave, current, or storm action. The properties of the water can also change between imaging passes, which in turn affects the imagery produced by the sonar.

The critical step in change detection is the alignment of the two images to be compared. For sonar images, the worst-case scenario is when the seafloor is bland, i.e., there are no objects such as rocks or debris in the scene [Sæbø et al. 2011]. This thesis presents research into the alignment of bland sonar images in order to quantify the accuracy with which it is possible to perform this alignment.

An introduction to sonar imaging techniques is given in Section 1.1. Section 1.2

follows this with a description of the difference between the system resolution and the image resolution of sonar imagery. The structure of this thesis is then outlined in Section 1.3 along with the assumed knowledge of the reader.

1.1 SONAR IMAGING

The name 'sonar' was originally derived from the expression **so**und **n**avigation **a**nd **r**anging. The first recorded mention of the use of sound for underwater detection is made by da Vinci in 1490 who wrote of the use of a tube inserted in the water through which the noise of other ships could be detected [Fahy 1998]. The first patent for an underwater echo ranging device was filed one month after the sinking of the *Titanic* and was followed by a number of others. Echo location research was driven by the need to counter the submarine threat during World War One, and has subsequently been developed into a wide range of sonar techniques. The reader is directed to works by Urick [1975] and Hunt [1978] for a more extensive history of sonar imaging.

Sonar is a time-of-flight technique: the transmitter sends out an acoustic signal and the receiver detects any echoes from objects in the scene. The delay between transmission and reception is a function of the distance to an object and the speed of propagation of the acoustic wave[1]. Each pulse is known as a *ping*. In this thesis, side-scan sonar is employed. This is where the sonar travels in one direction (known as the *along-track* direction) and images in the direction perpendicular to its motion (the *across-track* direction). Other geometries exist but are not considered here.

The attenuation of the transmitted signal is discussed in Section 1.1.1. Detecting the returned echoes is a key step of sonar imaging and is described in Section 1.1.2. The two main types of side-scan sonar, narrow-beam sonar and synthetic aperture sonar (SAS), are then detailed in Sections 1.1.3 and 1.1.4 respectively; this thesis exclusively uses SAS imagery. The raw data collected by a SAS needs to be processed or *reconstructed* to form the output image. A number of techniques for this image reconstruction are briefly outlined in Section 1.1.5.

1.1.1 Attenuation

The main loss in the system is the spreading loss, caused by the fact that a transmitted pulse will spread out as it moves away from the source. This results in a reduction of the energy density. For a spherically spreading wave, the energy density relative to the density at the source is given by $1/(4\pi r)$, where r is the distance from the source. In addition to the spreading loss, the signal loses energy from interactions with the medium it is moving through [Fisher and Simmons 1977]. Due to the density of water, acoustic (pressure) waves are able to travel with minimal attenuation at low frequencies.

[1] Typically $\sim 1500\,\mathrm{m/s}$ in sea water.

As the frequency increases, the attenuation of the transmitted signal increases [Urick 1975]. This attenuation is largely the result of friction, thermal conduction, and (in sea water) the chemical relaxation of the ions in the water [Fisher and Simmons 1977]. Typical losses are on the order of 0.1 dB/km at a frequency of 1 kHz, 1 dB/km at 10 kHz and 100 dB/km at 100 kHz. In terms of sub-seafloor imaging, the attenuation depends on the composition (particle size, density, etc.) of the sediment that the seafloor is made up from. This is a complex process, and the absorption loss it causes is generally obtained from published tables of empirical measurements [Hamilton 1980; Hamilton and Bachman 1982].

1.1.2 Echo detection

Having recorded the echoes returned from any objects in the scene, it is necessary to locate the ranges at which these objects are positioned. As with any real-world system, the recorded echoes will be corrupted by noise. The standard approach is to assume this is additive white Gaussian noise (AWGN). In this case the optimal detector is a matched filter. This is equivalent to correlating the received signal with a copy of the transmitted pulse [Crocker 1998]. Therefore the shape of the detected echo is governed by the autocorrelation of the transmitted waveform. The strength of the detected echoes relative to the noise is proportional to the amount of energy that was received; this is in turn a function of the amount of transmitted energy.

The *resolution* of the system defines its ability to distinguish two adjacent targets from each other. In terms of echo detection, this depends on the width of the autocorrelation: the narrower this is, the closer two targets can be before they cannot be individually detected. For a gated sinusoidal pulse of duration τ_p the autocorrelation has a triangular envelope of width $2\tau_p$. Hence, the resolution is directly proportional to the length of the pulse τ_p [Caprais and Guyonic 1997]. Reducing τ_p can improve the resolution, but requires an increase in the instantaneous power of the transmitter to maintain the total transmitted energy and thus the strength of the received echoes. However, the power is limited by non-linear effects such as cavitation [Urick 1975].

The standard approach is to increase the bandwidth of the signal by modulation. For a given bandwidth, B, the autocorrelation function is proportional to $\mathrm{sinc}(B\tau_p)$; the width of the main peak in the autocorrelation is then $1/B$ [Skolnik 1980]. The resulting resolution in the range (and, after image reconstruction, the across-track direction) is given by

$$\delta_r = \frac{c}{2B}, \tag{1.1}$$

where c is the speed of propagation of the signal. A commonly used signal is the linear frequency modulated (FM) chirp where the frequency of the signal linearly increases (or decreases) over a set range of frequencies. Other signals which have a narrow autocorrelation can be utilised, for example, pseudorandom sequences [Hawkins 1996].

It will be noted that the resolution (1.1) is not proportional to the pulse length, and so a low-power transmitter can use a long duration pulse to increase the total transmitted energy. The minimum range the system will measure (often the seafloor directly below the sonar) determines the maximum pulse length. Due to attenuation, the energy of the received echoes is several orders of magnitude lower than the transmitted energy. An echo returning while the transmitter is still running is liable to get swamped and thus ignored. Therefore, the pulse needs to have completed transmission prior to these close echoes returning.

1.1.3 Narrow-beam sonar

As the name implies, narrow-beam sonar uses a narrow beam of acoustic energy to insonify the scene. As shown in Figure 1.1a, the movement of the sonar between pings samples the scene. These are then combined to form a complete image of the scene. Ideally the *footprint* (the projection of the beam onto the seafloor) would have a constant width, but in reality it will widen as the range increases. The along-track resolution of a narrow-beam sonar can be approximated as [Caprais and Guyonic 1997]

$$\delta_y = \frac{r\lambda}{D}, \tag{1.2}$$

where r is the range, λ is the transmitted wavelength, and D is the width of the transmitter aperture in the along-track direction. As the range increases the resolution worsens, i.e., it is *range-dependent*. It can be improved by increasing the frequency of transmission (assuming the speed of propagation is constant, this decreases the wavelength), but this is limited by the increased attenuation of higher frequencies in salt water. Making the aperture longer also improves the resolution; again, this is limited by practical reasons, especially from a manufacturability point of view.

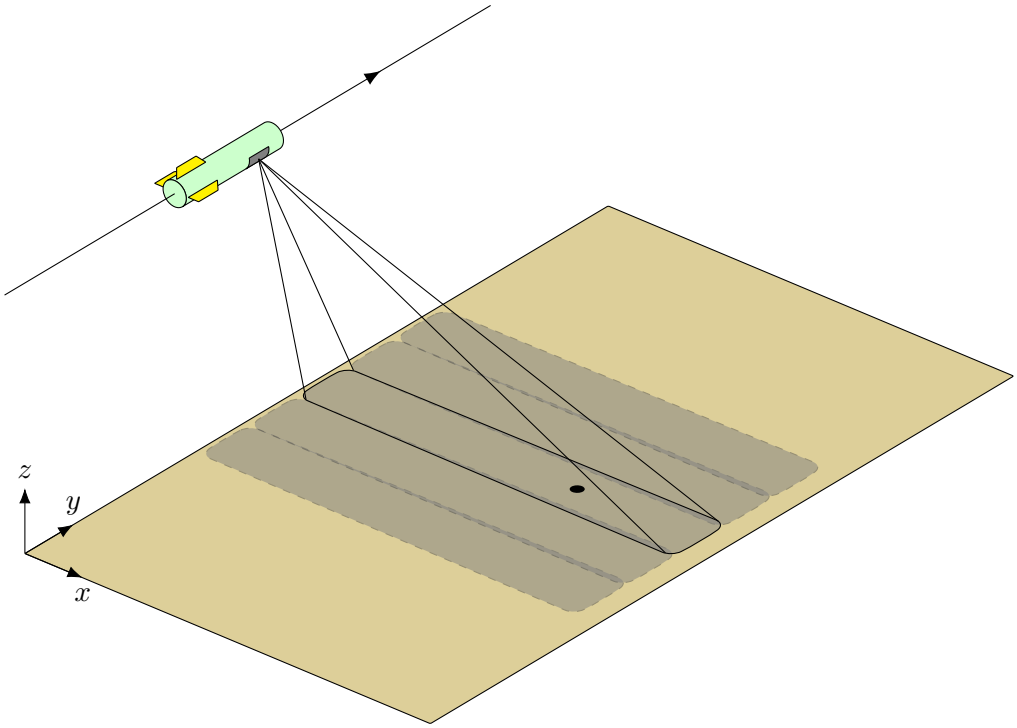
1.1.4 Synthetic aperture sonar

Synthetic aperture sonar takes the opposite approach to narrow-beam sonar. A wide beam is used to insonify the seafloor and, as illustrated in Figure 1.1b, this means that each target appears in a number of pings. The collected data can then be coherently processed to synthesise an aperture that is longer than the physical aperture of the transmitter [Curlander and McDonough 1991; Soumekh 1994], hence the name.

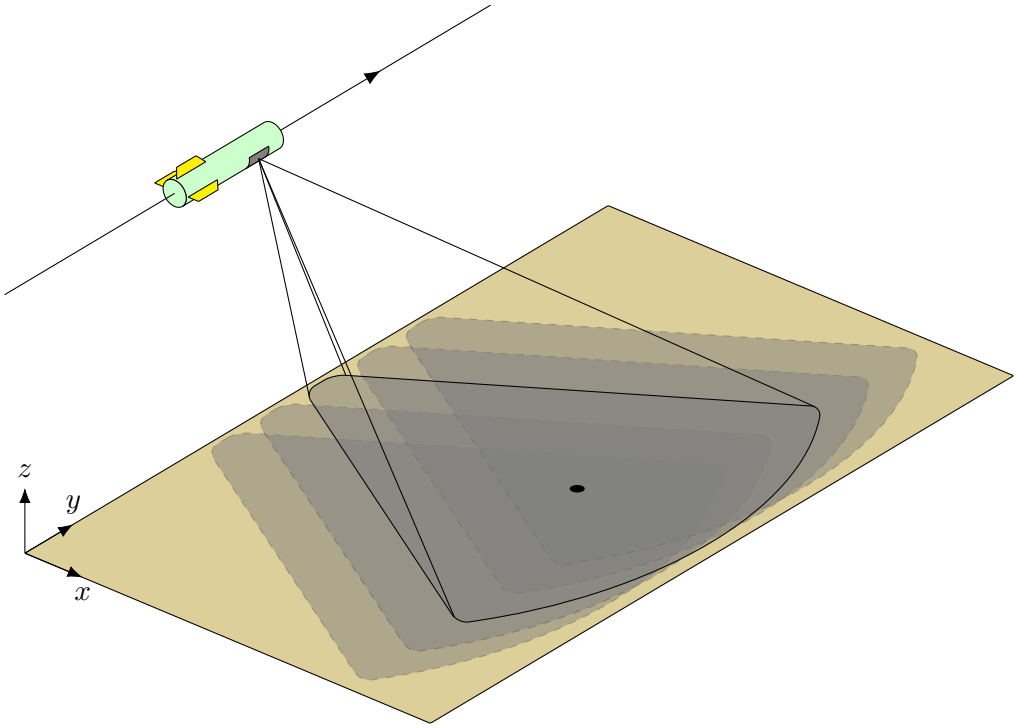
The along-track resolution of a SAS system is given by

$$\delta_y = \frac{r\lambda}{2L}, \tag{1.3}$$

where L is the length of the aperture that was synthesised during processing. This length depends on the range: the further away a point is, the more pings it appeared



(a) In narrow-beam side-scan sonar, each ping insonifies and images a small strip of the seafloor. The minimal overlap between the successive strips allows direct formation of the final imagery.



(b) With synthetic aperture sonar, a wide beam is used to insonify a large region of the seafloor. Each point of the seafloor appears in multiple pings, and coherent processing is required to form the final imagery.

Figure 1.1: A comparison of the geometries of (a) narrow-beam and (b) synthetic aperture sonar. As with any side-scan sonar, the platform moves in the along-track (y) direction, and images in the across-track (x) direction.

in. The length is given by

$$L = \frac{r}{\lambda D}. \tag{1.4}$$

Substituting this into (1.3) simplifies the expression for the along-track range to

$$\delta_y = \frac{D}{2}. \tag{1.5}$$

This is independent of both range and frequency. It is also inversely dependent on the length of the physical aperture. The smaller the physical aperture is, the wider the beam is, allowing a longer synthetic aperture to be formed.

1.1.5 Image reconstruction

The data collected by a SAS system requires processing to obtain an output image. This is the process of forming the synthetic aperture. A number of algorithms have been developed to perform this image reconstruction. A subset of these are described briefly in this section. The method used for the data presented in this thesis is the backprojection algorithm. For a more detailed background into SAS reconstruction techniques the reader is directed to works by Hawkins [1996] and Callow [2003].

The approach used in early SAS imaging was the correlation or matched filtering algorithm [Neilson 1991]. An expected point-spread function (dependent on both the range and along-track position) was defined for the scene. This was then correlated with the echo data to form the output image. Due to the computational effort required to perform this correlation, this algorithm — although exact — is inefficient. A variant of this, fast correlation, divides the scene into blocks and approximates the point-spread function at each point in a block by the point-spread function of the centre point [Hayes and Gough 1992]. The correlation is then performed in the Fourier domain.

Backprojection is a reconstruction technique used extensively in medical imaging for processing computed tomography (CT) data [Cho, Jones, and Singh 1993] and can be applied to SAS image reconstruction. The response from each ping is projected over a spherical arc to all the possible points in the image which contributed to the response. These projections are summed across all pings to generate the final image. Backprojection is more efficient than the correlation algorithm [Callow 2003] and shares the property that it can be used to reconstruct data obtained from an arbitrary track.

The wavenumber algorithm was first used in the seismic imaging field [Stolt 1978] and subsequently modified for use with synthetic aperture radar (SAR) [Cafforio, Pratti, and Rocca 1991] and SAS [Hawkins and Gough 1995]. The data is taken into the Fourier domain and the inverse Stolt mapping is applied to transform it from the measurement domain to the image domain.[2] This remapped data is then inverse Fourier transformed to yield the reconstructed image. The wavenumber algorithm is

[2]This reverses the mapping applied by the system during the collection process.

efficient but has the drawback that it requires regularly sampled data. For scenarios where the sonar did not travel in a straight line — or where it is desired to make local updates to the assumed path in order to align the image with one from a different trial — the wavenumber algorithm cannot be applied directly.

1.2 SYSTEM RESOLUTION AND IMAGE RESOLUTION

The system resolution is defined as the smallest individual feature that the sonar can resolve. This area is often referred to as a *resolution cell*. For a SAS system, the system resolution is defined by (1.1) in the across-track direction and (1.5) in the along-track direction. These are based on the physical properties of the sonar, the signal it transmits, and the medium it is operating in; it cannot be improved after the data has been collected.

The image resolution is the size of the pixels in the reconstructed images. This is independent of the system resolution, although it makes sense to base the image resolution on the system resolution. The relationship between the system and image resolution can be divided into three categories:

1. The image resolution is the same as the system resolution. In this case, each pixel in the image corresponds to a resolution cell of the sonar.

2. The image resolution is smaller than the system resolution. This means that each resolution cell is spread across multiple pixels in the image. Therefore the resolution cell has been interpolated or *oversampled* to generate the image. Oversampling may reveal finer-scale details of the scene, but it makes the assumption that the echo data is band-limited.

3. The image resolution is larger than the system resolution. Each pixel represents the average response of multiple resolution cells. Such an image is said to be *undersampled*. This effectively discards information collected by the sonar but may be useful for noise reduction or similar purposes.

1.3 THESIS OUTLINE AND ASSUMED KNOWLEDGE

Chapter 2 introduces the process of change detection with complex images and includes examples of this being performed on both radar and sonar imagery. The history of change detection in these fields is also given. SAS images are typically analysed statistically; the various statistical measures used for both individual images and pairs of images are presented in Chapter 3. Change detection requires the images to be precisely aligned, and developing algorithms to perform this alignment needs data with a known ground truth in order to test their accuracy. The data used in this thesis is

primarily simulated as this provides the necessary ground truth. Chapter 4 outlines the implementation of the SAS simulator used, along with some modifications that were required to generate valid repeat-pass data. Some simulations using these modifications are then presented in Chapter 5 validating the simulated results against the statistical models previously outlined. Chapter 6 then presents work on registration algorithms, including a study using high-quality navigation data as well as methods for estimating the alignment errors from the data itself. Finally, conclusions are given in Chapter 7 along with some suggestions for further research.

The preceding description of sonar imaging should be sufficient to understand the models and processes used within this thesis. The references given in those sections are available should the reader want a more in-depth discussion of sonar design, operation, or reconstruction algorithms. The main assumption made is that the reader has some familiarity with statistical signal and image processing. Chapter 3 provides details of the general techniques used subsequently in this thesis.

Chapter 2

CHANGE DETECTION

In a general sense, change detection is the process of comparing multiple recordings of an object or scene taken at different times in order to determine what attributes of that object or scene have changed in the interval. Historically, these changes were found from photographs by manual analysis. For example, in World War One, aerial photographs were used to detect changes to the enemy's trenches by projecting an image[1] onto a screen and tracing salient features onto the screen [Campbell 2008, p.87]. A later photograph could then be projected onto the same screen and, after careful alignment, any new features were readily visible. Similar manual techniques continued to be used throughout the 20th century for both military and civil applications [Théau 2008]. With the rise of digital processing in the 1970s, *automatic* or *digital* change detection started to replace manual methods [Lunetta and Elvidge 1998]. In addition to the applications with sonar and radar systems detailed later in this chapter, automated change detection has been used in fields as diverse as monitoring ecosystem [Coppin et al. 2004] and land cover changes [Mas 1999], medical imaging [Venot et al. 1984], driver assistance systems [Gerónimo et al. 2010; Kan, Krogmeier, and Doerschuk 1996], and monitoring civil infrastructure [Landis et al. 1999].

Section 2.1 discusses the use of change detection with complex datasets. A number of common change detection techniques are then outlined in Section 2.2. Finally, the history of change detection with SAR and SAS imagery is presented in Section 2.3.

2.1 CHANGE DETECTION WITH COMPLEX DATA

Imaging techniques such as SAS and SAR provide complex data, i.e., each pixel in the reconstructed imagery has both a magnitude and a phase associated with it. Change detection can be performed on the magnitude alone, in which case it is known as *incoherent change detection*. However, the phase of an image conveys more information about the image than the magnitude [Huang, Burnett, and Deczky 1975; Oppenheim and Lim 1981], and so discarding the phase could result in smaller changes being

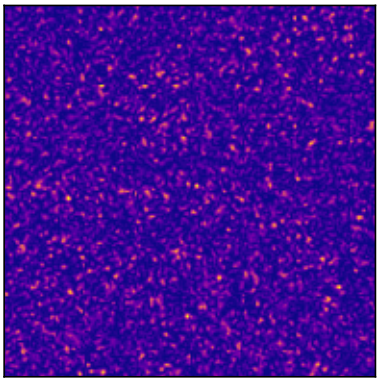
[1] Typically captured on glass photographic plates.

ignored. Utilising the phase as well as the magnitude leads to *coherent change detection*; coherence is a statistical measure of the degree of linear similarity between two datasets. It is introduced, along with its estimation over a spatial region and the factors which decrease it, in Sections 3.4 to 3.6. For the purposes of this chapter, it is sufficient to treat coherence as a number between 0 and 1 indicating the level of similarity between two images, with 1 meaning they are identical and 0 meaning there is no linear similarity.

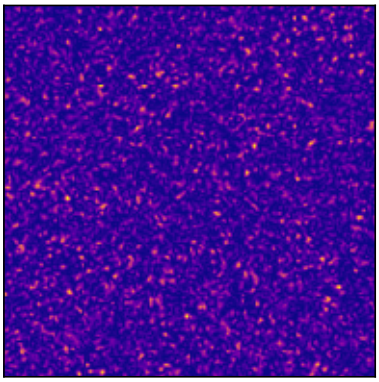
A simple example showing the sensitivity of coherent change detection is presented in Figure 2.1. The magnitudes of two speckle images, equivalent to the images that a SAS system would record from a bland seafloor[2], are given in Figure 2.1a and Figure 2.1b. These have been constructed such that the outer regions of the images are identical, but a circular region in the middle is statistically independent. This can be observed by careful inspection of the images, but it is not immediately obvious. One incoherent method of change detection is to look at the difference in the root mean square (RMS) power between corresponding points in each image. This is given in Figure 2.1c; the changed region is obvious in this case, but the magnitude of the change is low — the maximum change is 0.86 dB, and only 7 % of the changes are above 0.5 dB. The coherence between the two images is shown in Figure 2.1d. The unchanged region has the maximum possible coherence of 1, while the changes are close to 0. Figures 2.1e and 2.1f show the same incoherent and coherent change detection methods in the presence of additive white Gaussian noise. Although the 10 dB signal to noise ratio (SNR) used in these examples would be considered poor in a real-world situation, it serves to illustrate the sensitivity of the phase component: while the energy change has become harder to detect due to the changes in the background, the coherence map still clearly shows the change even though the magnitude in the background region has dropped.

Preiss and Stacy [2006] undertook a study of change detection with airborne SAR imaging. One of their experiments involved deliberate modifications to the imaged scene between imaging passes in order to demonstrate the effectiveness of coherent techniques. The resulting images are reproduced in Figure 2.2 (these correspond to Figures 47 to 50 in their report). Reconstructed images from the primary and repeat passes are shown in Figure 2.2a and Figure 2.2b respectively. The scene consists of a lightly grassed field bordered by trees and with drainage ditches running through it. Some buildings are also present amongst the trees at the left-hand side of the scene. During the two hour interval between passes a number of lines were made in the field using both a rotary hoe and a lawnmower. The (incoherent) change in the backscattered power is given in Figure 2.2c and the coherence between the passes is in Figure 2.2d. The modifications to the field are clearly visible as lines of low coherence but are not obvious when looking at the change in power. The trees also show up as changes due to the natural movement of the leaves and branches. In a pair of images of a different field in the same study, a number of irregular tracks through the middle of

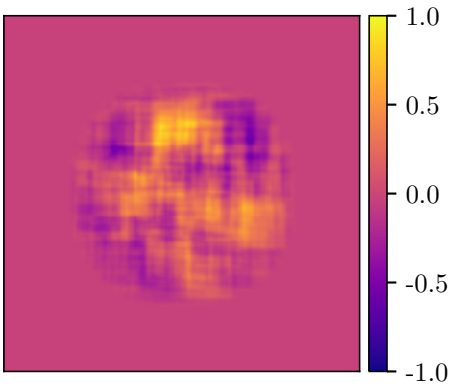
[2]See Sections 3.1 to 3.3 for details on speckle.



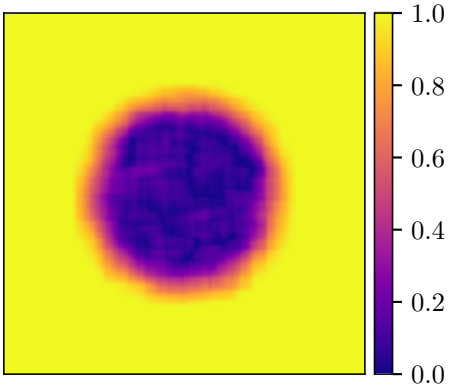
(a) First speckle image.



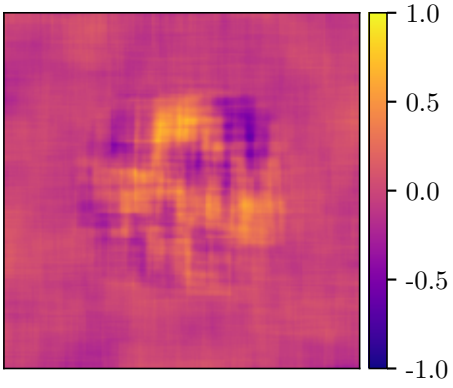
(b) Second speckle image.



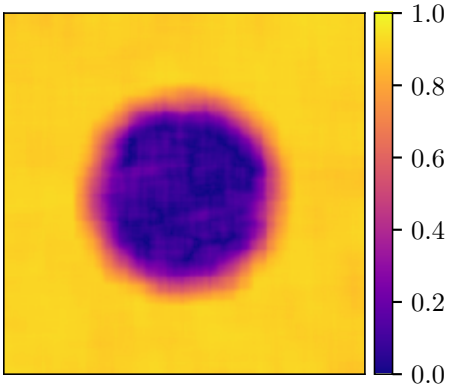
(c) Change in energy in decibels (no noise).



(d) Coherence (no noise).



(e) Change in energy in decibels (10 dB SNR).



(f) Coherence (10 dB SNR).

Figure 2.1: An example of coherent and incoherent change detection. Two speckle images are shown in (a) and (b); they are identical around the edges, but a circle in the centre is a different realisation of speckle. The difference in the energy between the two images is shown in (c), and the coherence between them in (d). The difference in energy is observable although the magnitude of the change is not significant. When white Gaussian noise is added to the images, the change in energy (e) is less obvious, while the coherence (f) still clearly shows the change.

the field were observed in the coherence. Although no definite ground truth is available to confirm it, sheep were subsequently observed in that field and therefore the observed tracks are likely to be caused by the sheep grazing.

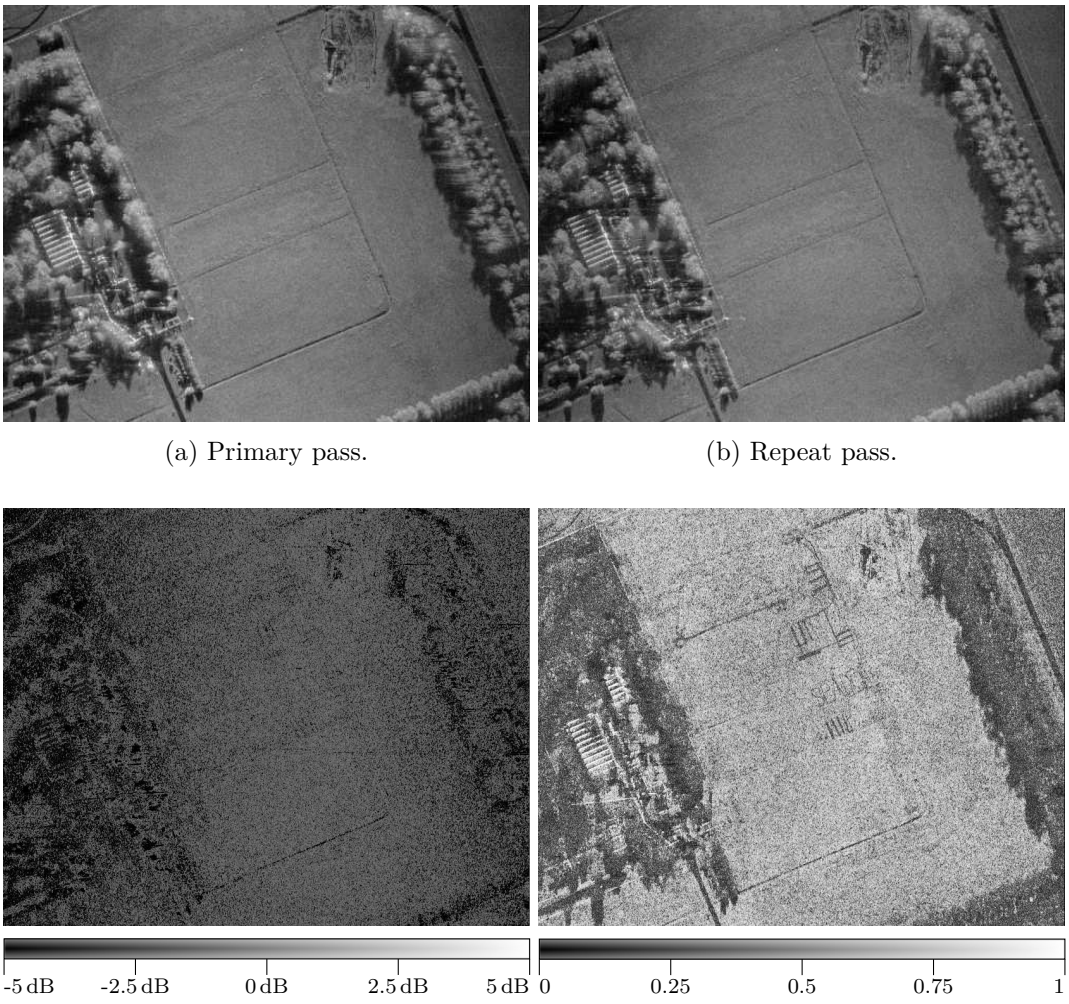
A good example of the sensitivity of coherent change detection with SAS imagery is given by G-Michael, Marchand, Tucker, Marston, et al. [2016] (Fig. 17 in their paper) and is reproduced in Figure 2.3. The primary pass over a sandy seafloor is shown in Figure 2.3a and a repeat pass several hours later is shown in Figure 2.3b. Performing incoherent change detection on these two images results in the difference map of Figure 2.3c. Small changes are observable in the top-left and bottom-right corners, but the majority of the scene is marked as having no changes. However, the coherence between the passes (Figure 2.3d) reveals a trench or track running between the two locations observed in the incoherent case. The authors surmise that this has been caused by some form of underwater creature. The other spots of low coherence around the image may also be caused by fauna, or may be a result of sand being shifted around by currents or wave action.

2.2 CHANGE DETECTION TECHNIQUES

There are many change detection techniques, and the most appropriate one for use depends on the application. They can be broadly divided into two categories: image-based techniques and feature-based techniques. The former directly compares the corresponding pixels in each image and applies some metric to show how much they have changed. Feature-based techniques first detect features or objects in a scene, and then compare the set of features found in each image to identify any differences.[3] This has the advantage of requiring less accurate registration or alignment of the images: if identical objects are found in similar locations, it can be assumed that they are the same and any position difference is due to an error in alignment. Image-based techniques require a highly accurate registration as comparing different pixels limits the usefulness of the result. The trade-off is that feature detectors require some *a priori* knowledge of the type of feature they are looking for whereas an image-based method can operate with the image alone.

This section outlines the common change detection techniques that have been used with sonar imagery in the published literature. For further details, including techniques used in other areas of research, the reader is directed to the various review papers on change detection techniques that have been published. These include works by Singh [1989], Lu et al. [2004], and Radke et al. [2005].

[3]In many remote sensing fields the feature-based approach is known as supervised detection as it operates on data that has previously been classified in some manner (e.g., into land use categories for satellite images of the earth). Unsupervised detection refers to image-based approaches which do not rely on information apart from what is present in the image itself [Bruzzone and Prieto 2000].



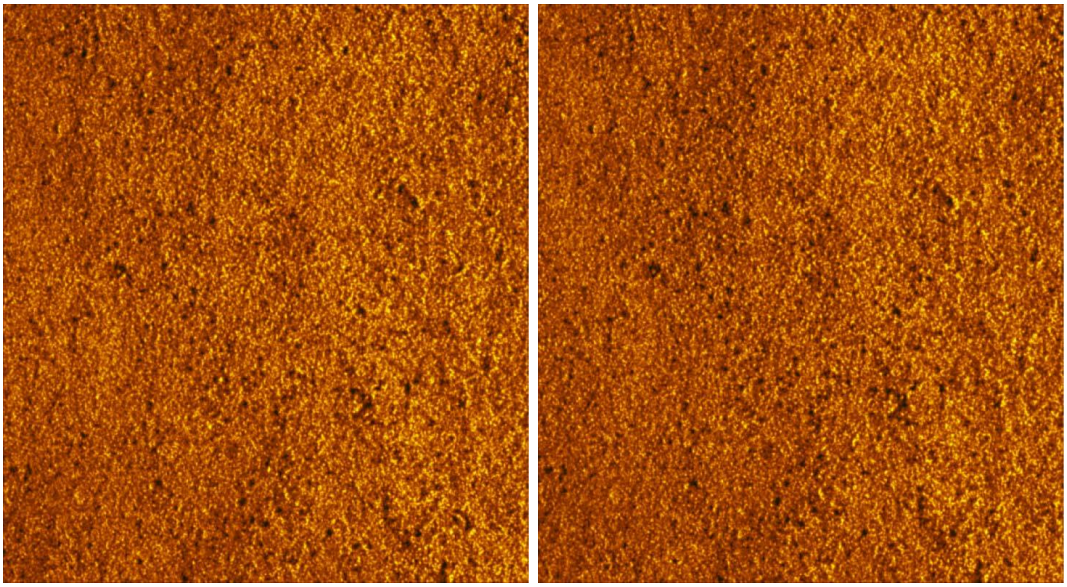
(a) Primary pass.

(b) Repeat pass.

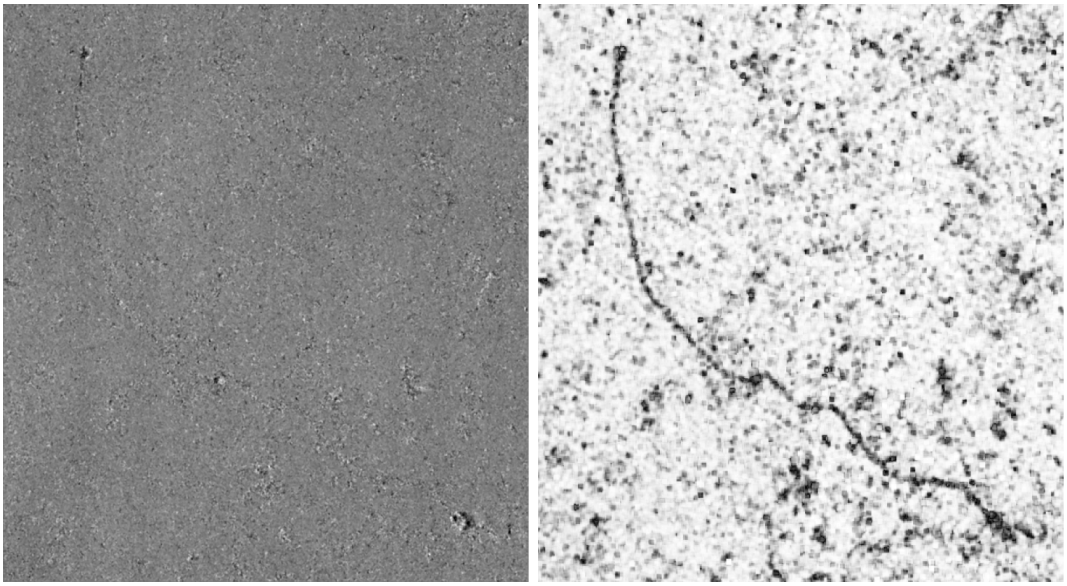
(c) Change in backscattered power.

(d) Coherence between passes.

Figure 2.2: A comparison between incoherent and coherent change detection from a SAR study by Preiss and Stacy [2006]. Intensity images of the primary and repeat passes are shown in (a) and (b) respectively. The passes were taken two hours apart, and the scene was modified by using a rotary hoe and lawnmower in the field during the interval. In (c) the change in the backscattered power (an incoherent statistic) shows no evidence of these changes. However, they are clearly visible in the coherence map shown in (d).



(a) Primary pass. (b) Repeat pass.



(c) Change in backscattered power. (d) Coherence between passes.

Figure 2.3: An example of change detection with SAS imagery from G-Michael, Marchand, Tucker, Marston, et al. [2016]. Two passes several hours apart over a sandy seafloor are shown in (a) and (b). Taking the logarithm of the intensities of these images and dividing gives the (incoherent) difference map shown in (c). Some small changes can be seen in the upper-left and lower-right corners, but the difference map is largely featureless. However, the coherence map in (d) shows a track or trench running across the scene, probably caused by fauna.

2.2.1 Feature-based techniques

The main application of sonar change detection to date — and practically the only one to use feature-based methods — is mine hunting. A standard method for finding mines is to look for objects which are standing proud of the seafloor. These can be identified by a bright spot (the target itself being assumed to be more reflective than the surrounding seafloor) and an adjacent shadow [Gendron, Lohrenz, Layne, et al. 2004]. The size and shape can be used to ignore detected objects which do not match the assumed profile of the targets of interest. Another technique is to have a set of template images representing either the objects of interest [Coiras et al. 2008] or the shadows they cast [Ferrand and Mandelert 2012]. Template matching, for example via correlation, can then be used to locate any such objects in the scene.

Having found a set of objects of interest in an image, the second step is to determine which ones represent changes. When comparing two images, this is a matter of calculating which objects appear in both; the ones that are missing from this union have been either added to or removed from the scene between the two passes. In a regularly surveyed area, such as a harbour or shipping lane, a historical database of objects may be kept. Objects detected in an image which do not appear in the database are thereby marked as changes, and can be marked for further checking or processing [Gendron and Lohrenz 2007].

2.2.2 Image-based techniques

For incoherent change detection, only the magnitude information of the imagery is used. It is standard to use the intensity or magnitude squared of the images, i.e., $I_1 = |d_1|^2$ and $I_2 = |d_2|^2$ for the two images d_1 and d_2. The two standard methods of change detection are *differencing*, where the intensities are subtracted pixel by pixel to give the difference $d = I_2 - I_1$, and *ratioing*, where the intensities are divided pixel by pixel to give the ratio $r = I_2/I_1$ [Rignot and Zyl 1993]. The ratio is often displayed in decibels, and in some uses of differencing the logarithm of the intensities are taken first in order to reduce the dynamic range. Note that these cases result in the two metrics being identical as

$$10\log\left(\frac{I_2}{I_1}\right) = 10\log(I_2) - 10\log(I_1). \tag{2.1}$$

In the examples of incoherent change detection presented previously (Figures 2.1c, 2.1e, 2.2c and 2.3c) ratioing was used. In these images — Figure 2.2c especially — grainy noise can be observed, and it is desirable to minimise this without affecting changes of interest. Often this reduction is performed by applying a moving average to the intensities to smooth out this noise. Another method that has been used with some success on sonar imagery is anisotropic diffusion filtering [Midtgaard et al. 2011] which is a technique designed to remove noise without affecting features such as edges or lines

in the image [Perona and Malik 1990]. A threshold can be used to ignore changes below a set level, typically chosen *a priori* based on the expected change for the target types of interest.

Statistical methods have also been investigated for incoherent change detection, although none have been widely adopted at this point. Myers, Fortin, and Simard [2009] applied three statistical measures: the relative entropy or Kullback-Leibler divergence [Kullback and Leibler 1951] between two sets of samples, the two-sample Kolmogorov-Smirnoff test [Massey 1951] for checking if the samples were from the same statistical distribution, and the Bray-Curtis distance [Bray and Curtis 1957] to measure the divergence in texture of the images. These detected the mine-like objects that the study was targeting, although they were outperformed by a feature-based technique. The salience (or prominence of some characteristic of the image) was used by Matthews and Sternlicht [2011] and shown to successfully detect targets in another mine-hunting experiment.

The prevalent method of coherent change detection is based around generating a coherence map of the scene. The images shown in Figures 2.1d, 2.1f, 2.2d and 2.3d are all examples of coherence maps. Each pixel in these maps gives an estimate of the coherence between the two source images at that point. As outlined in Section 3.5, the coherence is typically estimated over spatial windows surrounding the pixel of interest, and the estimate follows well-known statistical distributions. Having obtained the coherence map, further processing can be applied to highlight changes of interest. This may employ thresholding based on the statistics of coherence, algorithms which locate features such as lines, selecting low-coherence areas based on shape or size, or other similar techniques depending on the application.

Another method that has been used is canonical correlation analysis (CCA). This is a general statistical technique for investigating the relationship between two datasets first described by Hotelling [1936]. Given two vectors $\mathbf{x}$ and $\mathbf{y}$ containing observations of random variables, CCA involves finding vectors $\mathbf{a}$ and $\mathbf{b}$ such that the correlation coefficient[4] $\operatorname{corr}(\mathbf{ax}, \mathbf{by})$ is maximised. The first pair of canonical variables are then given by the products $\mathbf{ax}$ and $\mathbf{by}$. The second pair of canonical variables is found by repeating the maximisation with the added constraint that the second variables are independent from the first. This procedure can be iteratively applied to generate the set of canonical variables for the input vectors. These variables linearly map the data into their canonical coordinates; this is the coordinate system that is ideally suited for measuring the linear dependence between them [G-Michael, Marchand, Tucker, Sternlicht, et al. 2014]. As well as generating coherence maps [G-Michael, Marchand, Tucker, Marston, et al. 2016], CCA has been used to extract changed features from sonar imagery [G-Michael and Tucker 2010].

[4]See Section C.2 for the definition of correlation with random vectors.

It is worth noting that there is value in using both incoherent and coherent change detection techniques simultaneously as they can measure different changes. In general, incoherent techniques show the change in the amount of backscattered power measured by the receiver i.e., they indicate a change in the reflective properties of the area being imaged. On the other hand, coherent techniques detect changes in the structure of the target. A loss of coherence with no corresponding incoherent change indicates that the structure of the surface being imaged has changed (for example, it has been deformed) but that the material is still the same as it scatters the same amount of energy. Alternately, an incoherent change being observed with no loss of coherence suggests the object structure is unchanged but that the reflective properties of the material have altered. Rignot and Zyl [1993] observed such a situation with radar imagery, and postulated that this could be caused by a change in the moisture conditions which would change the dielectric properties of the ground — and thus the amount of reflected energy — without the surface shape changing.

2.3 CHANGE DETECTION HISTORY

This section details the history of change detection with both radar and sonar imagery. In the former case it is intended to be representative rather than comprehensive, while the sonar section is intended to be complete. As outlined in Section 2.2, in general change detection can be divided into image-based and feature-based approaches. For radar imagery, the early work focused on image-based techniques and feature-based work is relatively recent [Gamba, Dell'Acqua, and Lisini 2006]. The availability of accurate navigation data (especially with satellite-based radar) and the fact that the background is generally relatively stable (trees, mountains etc. do not tend to change much in the short or medium term) makes this approach possible. With sonar data, feature-based approaches form the bulk of the initial work as the navigational challenges and instability of the background have made image-based change detection infeasible until recently [Sternlicht, G-Michael, and Matthews 2012]. The fact that much of the early research was performed with sidescan sonar imagery is also a factor in this: having a range-dependent resolution and typically being amplitude-only makes it less suitable for image-based techniques compared to SAS [Sternlicht, Harbaugh, and Nelson 2009].

2.3.1 Change detection in radar

A major application in the early years was measuring the drift and deformation of sea or pack ice over time. Some initial work was done with optical images from the Landsat satellite for both near-shore [Hibler, Ackley, et al. 1974; Shapiro and Burns 1973] and off-shore [Hibler, Tucker, and Weeks 1975] regions. In the latter case, the absence of land-based (fixed) reference points meant the accuracy of the measurements

was dependent on the deviation between the reported centre of a given image and the actual position; errors up to 8 km were reported, although the average error was less than a typical days worth of drift. Leberl, Bryan, et al. [1979] used airborne SAR images taken sixteen days apart during the 1975 Arctic Ice Dynamics Joint Experiment of an area near Prudhoe, Alaska to perform the same task. As before, the accuracy was limited by the navigation data available: inertial navigation alone resulted in errors of up to 2.5 km, but this was able to be reduced to fractions of a kilometre if precisely-surveyed ground stations were present in the imagery. The relative drift between ice floes was determined to ±200 m, limited only by the radar geometry and ability to identify matching points between the images.

Launched into Earth orbit on June 27 1978, Seasat carried the first spaceborne SAR during its 105 days of operation[5] [Michalopoulos 1979]. Leberl, Raggam, et al. [1983] used images from seven passes over Banks Island, Canada to track the movement of sea ice, achieving an accuracy of ±500 m utilising known ground points for alignment. This work was extended by Fily and Rothrock [1987] who used a hierarchical approach, applying correlation over increasingly smaller regions of the image to achieve an accuracy of 75 m. Bryan and Clark [1984] examined Seasat images from repeat passes over a number of environments (urban, rural, desert, and wetlands), showing that change detection with Seasat SAR imagery was viable. For example, the difference between four passes over a rural area of Kansas exhibited noticeable changes in intensity due to moisture variations and agricultural practices. It was noted that for oriented features (streets, buildings, agricultural fields, etc.) a difference in the look angle from the radar between passes does cause a difference in the intensity. A number of space shuttle missions included a SAR in the payload. Although the main purpose of the data collected on these missions was the generation of an accurate elevation model of the surveyed area, some portions of the data have subsequently been used for change detection. One example is the use of two passes of the shuttle Endeavour one day apart to measure the flow velocity of Glaciar Moreno in Argentina [Michel and Rignot 1999].

A number of earth sensing satellites carrying SARs were launched in the first half of the 1990s. These include two European Remote Sensing satellites (ERS-1 in July 1991 and ERS-2 in April 1995), the Japanese Earth Resources Satellite 1 (JERS-1) in February 1992, and the Canadian RADARSAT-1 in November 1995. This lead to a large amount of repeat-pass SAR imagery becoming available. Rignot and Zyl [1993] used data obtained during the ERS-1 testing phase — when the orbits were very stable, and so the need for geometric corrections was eliminated — to evaluate its use in change detection. Both incoherent and coherent methods were tested, and were shown to be complementary techniques measuring different sorts of changes. A number of studies into the coherence of a variety of terrain when imaged with ERS-1 were also carried out [Corr 1997; Corr and Rodrigues 1999; Corr, Whitehouse, et al. 1996].

[5]The mission ended on October 10 1978 due to a short-circuit in the electrical system.

ERS-1 and ERS-2 followed the same orbit with ERS-2 one day behind ERS-1. Corr and Rodrigues [1998] used this to collect images one day apart over an area of Salisbury Plain in England which is used as an army training area. The terrain is predominantly grass, and on image pairs taken when the training area was in use, coherent change detection was able to observe tracks made by vehicles. ESR imagery was also used to detect the presence of planes and vehicles at Heathrow airport [Oliver, McConnell, and Corr 1999]. Liu et al. [2001] used coherence images derived from three ERS-1 passes (giving 35, 350, and 385 day temporal separation options) over an area of the Sahara desert in Algeria to detect a number of features. The movement of sand was observed, and from this the boundaries of sand dunes were determined. A number of ephemeral lakes[6] were identified along with the drainage and erosion patterns in the observed area. Finally, a number of seismic survey lines — used for oil exploration — were noticed in the coherence images.

Much of the recent work has continued this vein of using spaceborne SAR imagery for monitoring changes. For example, Chi, Sun, and Ling [2009] used data from a variety of satellite sources to show their potential in monitoring changes in an urban environment using Fuzhou city in China as a test site. Engeset et al. [2002] used satellite SAR data acquired over multiple years to monitor the mass of glaciers in the Svalbard region. There has also been ongoing research into improving the methods used to detect changes.

2.3.2 Change detection in sonar

Early change detection with sonar images took the form of operator aids rather than automatic detection. For example, the "blink" operator aid detailed by Poeckert [1991] was used to assist with mine detection. It rapidly alternated between displaying two sonar images on a CRT screen, one historic and one current. Differences between the images were then detected by the user as an object blinking. This system was based on the blink comparator designed for use with astronomical photographs by Carl Pulfrich in 1904 [Wolfschmidt 1998][7].

Bellec et al. [2005] conducted a feasibility study into repeat-pass interferometry[8] using low-frequency SAS data collected as part of a mine-hunting experiment. They outlined a number of factors that would impact the achievable coherence and showed that although the measured coherence was lower than their predicted value, interferometry was viable with repeat-pass data. The main note from the experiment was the need for accurate track information. A similar feasibility study into multipass coherent

[6] Lakes that are usually dry but temporarily fill with water after rainfall.

[7] The blink comparator was used in the discovery of many astronomical bodies, most notably Clyde Tombaugh's discovery of Pluto in 1930 [Croswell 1997].

[8] Comparing two images to estimate the *bathymetry* or height of a scene. This has a lot in common with change detection.

processing is presented by Synnes et al. [2010]. A SAS mounted on an autonomous underwater vehicle (AUV) was used to collect data in two experiments, each of which used two tracks that intersected at a given angle (5° in one experiment and 0.5° in the other). This was used to examine the level of coherence loss due to separation between the passes. They also showed that the navigation system of the AUV could be updated from the results of the coherent processing thereby improving the overall performance of the system.

In general, SAS interferometry has been performed in a single pass with multiple vertically-separated hydrophones, although a few subsequent works have looked at the repeat-pass case. For example, De Paulis et al. [2011] present a study of repeat-pass interferometry for monitoring seabed deformation over a period of up to two years. A set of trials carried out in the Bedford Basin off Nova Scotia was used by Dillon and Myers [2014] to demonstrate the technique. Image-based alignment procedures were used to yield an improved coherence between passes, and therefore more accurate interferometry. Hunter, Dugelay, and Fox [2016] present an algorithm which estimates errors in the relative paths taken by the sonar in the two passes. Knowledge of these errors allows the alignment of the images to be improved; this is demonstrated by the authors for a set of trials made with deliberate path differences.

Much of the early research into automated change detection was focused on searching for mines or improvised explosive devices (IEDs). The typical method of operation is to detect mine-like objects and compare them to a list of previously detected objects to identify any new targets. Gendron, Lohrenz, Layne, et al. [2004] introduced the Automatic Change Detection and Classification (ACDC) system for use with sidescan sonar. The detector operated by finding adjoining bright spots and shadows, working on the assumption that any object standing proud of the seafloor will cast a shadow. These were classified by comparison with a database of expected sizes and shapes of targets of interest, and a historical database of known objects was then used to discard any previously identified targets. This was subsequently extended to real-time operation [Gendron and Lohrenz 2007]. A later mine-hunting experiment in the port of Corpus Christi, Texas was used to test the ACDC system with SAS data [Gendron, Lohrenz, and Dubberley 2009]. Some months prior to the experiment the area had been imaged with sidescan sonars, and the same sidescan sonars were used during the experiment as well as the SAS system. The historical and current sidescan imagery was successfully used for object detection. However, the SAS system required downsampling and pre-processing to make its images emulate those of the historical sidescan sonar. This reduced the performance of the ACDC system although it still showed that the technique was feasible.

A similar approach was followed by Coiras et al. [2008]. Objects of interest were detected via correlating the image with a template of an idealised target. This was followed by data association being used to match and align targets with historical

imagery; targets which could not be matched indicated new objects. In order to reduce error rates, the authors showed that using multiple imaging runs over an area could feed into a persistence ratio: how many times an object was flagged against how many times the area was imaged, with a high ratio indicating more certainty about the target. It was also suggested that target recognition algorithms could be applied to discard types of targets that were not of interest.

Myers, Fortin, and Simard [2009] studied three statistical methods which considered the distribution of pixels in corresponding regions of two SAS images, and compared this to the output of a automatic target recognition algorithm. It was concluded that automatic recognition performed better (more changes detected, and under half the number of false alarms of the best statistical method). The statistical methods all detected the same changes but the areas they falsely marked as having changed differed from method to method; the authors suggest combining the output of these methods would improve the results, albeit at the cost of greater computational effort. Another statistical method, temporal invariant saliency, was introduced by Matthews and Sternlicht [2011]. After filtering to remove high-frequency noise, a local variance measure (the difference between the variance in a given local area and the variance of the entire scene) is used to isolate changes from the background. This technique has been used for mine hunting in harbours and high-volume traffic lanes [Sternlicht, G-Michael, and Matthews 2012].

In April 2011, a set of SAS trials — known as the Larvik trials — were carried out near the Oslofjord in Norway. A number of research institutes from different countries took part, and one of the aims was to collect data for change detection experiments. Two seafloor areas, one relatively bland and one rough and cluttered, were selected for the trials. The first scene was imaged with a variety of targets placed in it; the targets were then removed by divers and the scene re-imaged. The reverse was carried out with the second scene: it was imaged, the targets were added, and it was imaged again. Midtgaard et al. [2011] presented some preliminary incoherent change detection results, showing that, after appropriate filtering, the deployed objects were clearly visible in a difference image between the magnitudes of two passes. Quidu et al. [2012] looked at both incoherent and coherent change detection and found that, although coherent techniques worked well, for large targets as were deployed in the trials incoherent methods were sufficient. Extending this, Myers, Quidu, et al. [2013] showed that the coherent method detected smaller changes not observed by the incoherent method, for example, some drag marks left by divers removing the targets from the first scene. Their analysis of the second (rough seafloor) area showed that the coherence decayed more quickly than for the smooth seafloor, and so fine changes were not able to be detected as reliably.

A number of relatively recent publications have used image-based change detection as opposed to feature-based detection. Sternlicht, Harbaugh, and Nelson [2009] added

a simulated contact to the second image of a pair and showed that thresholding the resulting coherence found the change. This was then applied to a second set of data where a target was added to the seafloor before the second pass was performed. Canonical correlation analysis was used by G-Michael and Tucker [2010] to perform coherent change detection between two passes with targets added to the scene between the runs. Sæbø et al. [2011] used images collected from two passes taken 13 minutes apart over a flat sandy seafloor. Although no deliberate changes were made to the scene between passes, after the images were aligned a drop in the coherence was noticed in one small patch of the seafloor. The authors assume that this was due to fauna activity. A similar result was found by G-Michael, Marchand, Tucker, Marston, et al. [2016] and Sternlicht, Fernandez, et al. [2016]: after the two passes were aligned, a coherence map showed a trench presumably created by fauna[9].

The idea of combining imagery from multiple sonars was explored by Tucker and Azimi-Sadjadi [2011]. In change detection applications, clutter consists of objects which, although not actually targets of interest, appear like targets to detection algorithms. Low-resolution broadband sonar systems have better clutter suppression abilities than their high-resolution counterparts. Combining such a system with a high-resolution sonar for good target definition was shown to result in a significant reduction in the false alarm rate of the detector compared to using either system individually.

Midtgaard [2013] addressed change detection over long time intervals (up to three years) under the theory that, due to the cost of SAS surveys, many applications would require working with infrequent imaging. Experiments were carried out in Norwegian coastal waters using two sites, one muddy area left untouched and one sandy area with different targets added before each subsequent run. All change detection was performed incoherently. In the muddy site, there were some residual changes due to the paths of the sonar (and thus the position of any shadows cast on the seafloor) being different between runs, but no significant changes otherwise. With the sandy site, there was some background noise due to sediment transport, movement of shellfish etc., but the deployed targets were clearly distinguishable. The conclusion is that incoherent change detection is feasible for long time scales in suitable conditions.

Most of the previously-mentioned research considers only pairs of images when looking for changes. For multiple passes over a scene G-Michael, Tucker, and Roberts [2016] define a measure called the statistically normalised coherence (SNC). For a given pixel in a pair of images, it measures the deviation of the coherence compared to the mean coherence over all previously-seen pairs of images. Negative values indicate a lower coherence than average while positive values indicate a higher than average coherence. Application of this method to experimental data showed that using images from multiple passes reduced the problems caused by shadows and other artefacts when

[9]This is the example result shown in Figure 2.3.

detecting changes.

A relatively recent approach is to use feature-matching techniques to align the images. Developed in the computer vision community for use with optical images, methods such as the scale-invariant feature transform (SIFT) [Lowe 1999] and speeded up robust features (SURF) [Bay, Tuytelaars, and Van Gool 2006] operate by finding local features (portions of an image which stand out from their surroundings). These features are represented by a *descriptor* which encodes the properties of the feature, typically in a rotation- and scale-invariant manner. The features identified in each of a pair of images are then compared to find the *correspondences* or matching features. These correspondences can then be used to determine the difference in alignment between the images. For SAS images with a number of objects on the seafloor, both SURF [Midtgaard et al. 2011; Quidu et al. 2012] and SIFT [G-Michael, Marchand, Tucker, Marston, et al. 2016; G-Michael, Marchand, Tucker, Sternlicht, et al. 2014] have been shown to be practical for image alignment. Wang and Hayes [2014, 2016a,b, 2017] investigated the use of both algorithms on simulated bland seafloor images, and show that for a coherence lower than 0.9 feature matching becomes implausible due to a lack of common features. They present a simple two-parameter model which predicts the number of matching features as a function of the coherence and suggest that this acts as a lower bound for non-bland images (i.e., images which have objects lying on the seafloor).

Not all detected changes are the result of differences on the seafloor. Analysing some trials off Elba Island, Italy,[10] Hansen et al. [2015] noticed some large linear structures — tens of metres long by several metres wide — appeared in the initial pass and had disappeared or changed when the repeat pass was performed 28 hours later. The sea state was low and the changing structures could not be explained by sediment transport; this is reinforced by the fact the rest of the seafloor was largely unchanged. Noting that the changes were at a depth where there was a sharp change in the density of and speed of sound in water, the authors suggest the features were caused by a breaking internal wave[11] causing refractive effects within the water column. Simple geometric modelling of the situation fits these observations. Although this potentially interferes with seabed change detection, it does also suggest that modern SAS is capable of observing oceanographic properties which cause variations in the water column.

2.4 DISCUSSION

Change detection is, in general, a well established area of research. With sonar data it has been shown to have uses in mine hunting applications, and recent results with image-based coherent detection have been able to detect small changes made by fauna.

[10] Probably most well known as the island to which Napoleon was first exiled.

[11] A wave contained within a fluid medium rather than on its surface.

The key to enabling change detection is the registration of the images. Once the two sets of imagery are accurately aligned, there are a number of algorithms which can be used to identify areas of difference. The remainder of this thesis will focus on performing this alignment with bland sonar images.

Chapter 3

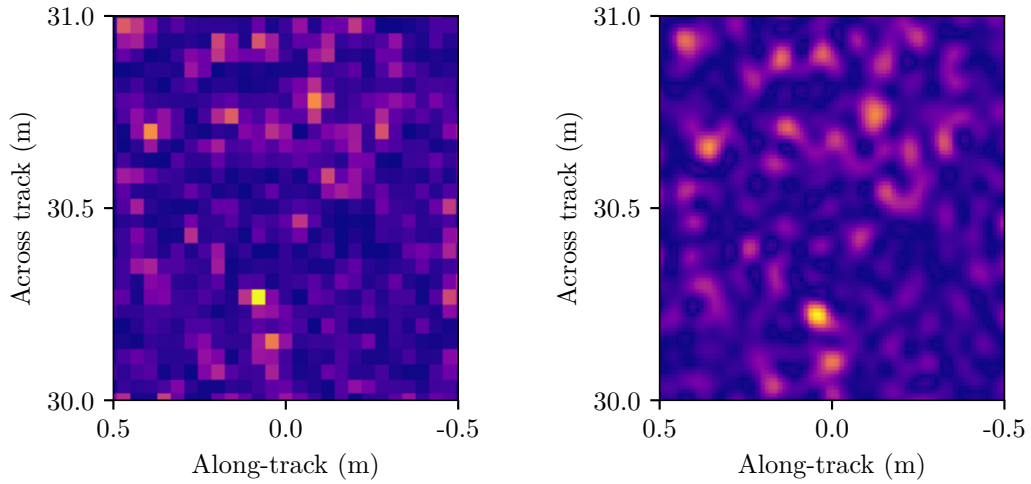
STATISTICS OF SAS IMAGERY

Many components of a SAS system can be considered as random variables: there are errors in the measured position and orientation of the system, the recorded echoes are subject to acoustic and electrical noise, the exact nature and position of the seafloor is unknown, and so forth. Therefore, it makes sense to also treat the reconstructed imagery as a random variable and to describe and analyse it statistically. This chapter details the statistics of SAS images along with their analysis as it pertains to change detection applications.

Due to its coherent nature, all SAS imagery contains *speckle*, a granular modulation caused by sub-resolution scatterers. Section 3.1 provides a description of how it occurs and its history in coherent imaging. Section 3.2 then details statistical modelling of speckle, and is followed by some metrics for speckle analysis and how they can be applied in Section 3.3. Coherence, an ensemble statistic measuring the degree of linear similarity between two signals, is introduced in Section 3.4. Coherence cannot be directly measured from a SAS image but instead has to be spatially estimated within some portion of the image. This is outlined in Section 3.5 along with various statistical models describing the accuracy of this estimation. Section 3.6 then details various factors, both within the sonar system and the environment, which degrade the coherence. The related concept of correlation is introduced and compared to coherence in Section 3.7. As all of the statistical measures presented in this chapter require repeated evaluation over sliding windows, Section 3.8 finishes by presenting an efficient technique of performing this evaluation using integral images.

3.1 SPECKLE

Speckle is a phenomenon present in all coherent imaging systems [Dainty 1976]. If a surface is rough at a sub-wavelength scale, then each resolution cell illuminated by the source will have multiple independent surfaces within it. The reflections from each of these constructively and destructively interfere to form the overall response. This results in the granular appearance commonly associated with speckle. Speckle is



(a) No oversampling.

(b) 4 times oversampled.

Figure 3.1: SAS intensity images from a simulated bland seafloor. The granular appearance is speckle. In (a), the image has been reconstructed with no oversampling i.e., the pixels are the same size as the sonar resolution cell, and so each pixel is an independent speckle. In (b), the image has been oversampled 4 times so that each pixel is one-quarter the size of the resolution cell. This does not affect the physical size of the speckle, but it means that each speckle is spread across multiple pixels.

also present in astronomical imaging where the fluctuations in refractive index of the earth's atmosphere causes short exposure images to exhibit the same granular appearance [Fried 1966]. It is worth noting that despite its random appearance, speckle is a deterministic phenomenon: if the same system is used to image a cell repeatedly from a fixed location, the sub-resolution scatterers will interfere in the same manner and thus an identical response will be observed. An example of speckle images from a SAS simulation of a bland seafloor are given in Figure 3.1. If the data is reconstructed so that the size of the pixels in the output image is the same as the size of the resolution cell of the sonar, then the image in Figure 3.1a is obtained. Each pixel in this image is an independent speckle. If the image is then oversampled four times in each direction, the image in Figure 3.1b is the result. Here, each pixel is one-quarter the size of the resolution cell in each direction (i.e., one-sixteenth the area) and is no longer independent from its neighbours. Note that this does not affect the physical size of the speckles; it just means that the speckle is spread across multiple (smaller) pixels. In general, the size of the speckle is inversely proportional to the bandwidth of the imaging system [Dainty 1976].

One of the earliest published observations of speckle was from Austrian physicist Karl Exner [1878, 1880] who reported on speckle patterns observed on a fogged glass plate illuminated by candlelight. Analysis by Lord Rayleigh [1880, 1918, 1919] provided a mathematical basis for the description of speckle patterns while a number of

contemporary papers discussed the radial structure present in the patterns; ultimately this was found to be due to the non-monochromaticity of the light [Dainty 1976]. The distribution of speckle intensity proposed by Rayleigh was subsequently quantitatively verified in a study of coronae [Ramachandran 1943]. For a more detailed bibliography of this initial research into speckle, the reader is directed to the works of Hariharan [1972] and Dainty [1976]. With the advent of laser imaging speckle became studied again. It was first observed in optical laser experiments published concurrently by Rigden and Gordon [1962] and Oliver [1963]. Works by Dainty [1975] and Goodman [1976, 1986] provide good references on the background and analysis of speckle in coherent imaging.

Initially speckle was viewed solely as a drawback due to its degradation of the observed data. Although it is still considered detrimental in many fields, the fact that speckle patterns contain information about the structure of the surface has been utilised in others. For example, electronic speckle pattern interferometry (ESPI) is capable of measuring surface displacement to a sub-micrometre level [Løkberg 1980] with applications including vibration analysis, flaw recognition, and the measurement of mechanical and thermal distortions [Sharp 1989]. With this technique a laser is used to illuminate a surface (which must be rough at optical wavelengths), and the resulting scattered light field is superimposed with a reference beam derived from the same source. The interference between the two light fields forms a speckle image. If the surface is then deformed, this speckle pattern will be altered with the change in the pattern being proportional to the amount of deformation that occurred [Petzing and Tyrer 1998]. The temporal resolution of ESPI is only limited by the camera and the power of the laser, with Moore et al. [1999] demonstrating framerates of 40,500 frames per second. Other applications of speckle analysis have been as varied as evaluating the quality of oranges after storage [Rabelo et al. 2005] and using the response of the human eye to a rotating speckle pattern to determine long- and short-sightedness as well as the presence of astigmatism [Ennos 1996].

3.2 SPECKLE STATISTICS

Due to incomplete information about the sub-wavelength structure and reflective properties of the scene, speckle can be modelled statistically [Kuttikkad and Chellappa 2000]. A common model is that of a circular Gaussian multiplicative noise process; this is known as *fully developed speckle*. Goodman [1986] gives a number of conditions required for fully developed speckle:

- Each resolution cell contains a large number of scatterers. If this is not true (i.e., if the size of each speckle is comparable with the resolution of the system) then analytical expressions for the speckle intensity distribution are not available. In

the limiting case of a system which resolves every scatterer, no speckle will be observed.

- Each scatterer is independent.
- The phase of each scatterer is uniformly distributed on $[0, 2\pi)$. If this is not the case, the assumption of circular Gaussian statistics is violated.

If these conditions are met, the speckle can be treated as the result of a random walk in the complex plane, with each step in the walk corresponding to the response from an individual scatterer. Applying the central limit theorem, the real and imaginary parts of the speckle field, denoted by random variables X and Y respectively, are uncorrelated, zero-mean Gaussian variables with identical variance σ_x^2 [Goodman 1975]. For SAS imaging, σ_x is proportional to the reflectivity of the seafloor. Therefore the joint PDF of X and Y is the uncorrelated zero-mean bivariate normal distribution,

$$f_{XY}(x, y) = \frac{1}{2\pi\sigma_x^2} \exp\left(-\frac{x^2 + y^2}{2\sigma_x^2}\right). \tag{3.1}$$

The speckle can be represented by combining the two components to yield a complex normal random variable $Z = X + jY$. A general complex normal distribution is characterised by three parameters: the location vector μ, the covariance matrix $\mathbf{K} = \mathrm{E}\big[(Z-\mu)(Z-\mu)^{\mathrm{H}}\big]$, and the relation matrix[1] $\mathbf{M} = \mathrm{E}\big[(Z-\mu)(Z-\mu)^{\mathrm{T}}\big]$. For two independent and identically distributed (i.i.d.) zero-mean variables such as X and Y, $\mu = 0$ and $\mathbf{M} = 0$, and therefore the combined distribution Z is circularly symmetric [Gallager 2008]. Circular symmetry means that the distribution is invariant to rotation in the complex plane, i.e., $e^{j\phi}Z = Z$ for all real ϕ, and follows from the requirement that the phase of each scatterer be uniformly distributed.

It is common to analyse the magnitude and the phase of SAS images independently. The magnitude is generally proportional to the backscattered energy from each region of the scene, while the phase can be used with techniques such as interferometry to recover fine details of the scene. The speckle magnitude $|Z| = \sqrt{X^2 + Y^2}$ is a Rayleigh-distributed variable,

$$f_{|Z|}(z) = \begin{cases} \dfrac{z}{\sigma_x^2} \exp\left(-\dfrac{z^2}{2\sigma_x^2}\right) & z \geq 0, \\ 0 & \text{otherwise.} \end{cases} \tag{3.2}$$

Figure 3.2a shows some examples of the PDF of the speckle magnitude for different variances. The image intensity — the square of the magnitude — is also used when processing SAS data. For speckle, the intensity $I = |Z|^2 = X^2 + Y^2$ follows a negative

[1] Also referred to as the *pseudocovariance* matrix in some literature.

exponential distribution:

$$f_I(i) = \begin{cases} \frac{1}{\sigma_i} \exp\left(-\frac{i}{\sigma_i}\right) & i \geq 0, \\ 0 & \text{otherwise}, \end{cases} \tag{3.3}$$

where $\sigma_i = 2\sigma_x^2$. A few examples of this distribution are given in Figure 3.2b. The n^{th} moment of the intensity is given by

$$\mathrm{E}[I^n] = n!\,\sigma_i^n, \tag{3.4}$$

allowing the mean of the intensity to be calculated as σ_i and the variance as σ_i^2. The speckle phase $\Theta = \arctan\left(\frac{Y}{X}\right)$ is uniformly distributed, i.e.,

$$f_\Theta(\theta) = \begin{cases} \frac{1}{2\pi} & -\pi < \theta \leq \pi, \\ 0 & \text{otherwise}. \end{cases} \tag{3.5}$$

This is a by-product of the circular symmetry of Z: since the distribution is invariant to rotation, it is also independent of angle.

3.3 SPECKLE ANALYSIS

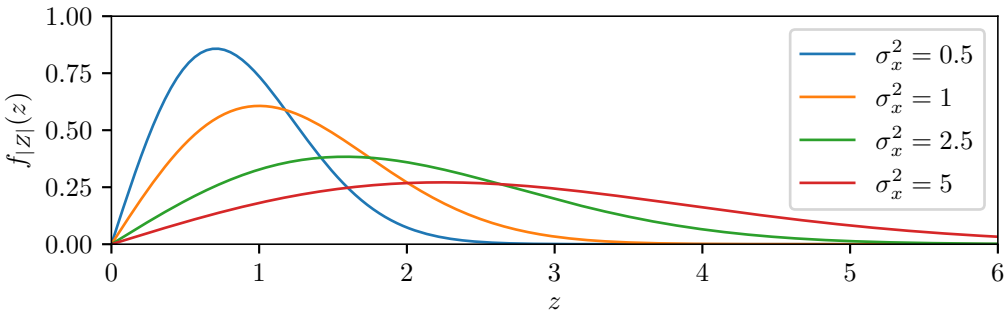
Measuring the level of speckle is a common stage in a SAS processing pipeline. Changes in the distribution of the measured intensity indicate a change in the scene, whether from a variation in the seafloor (differing sediment sizes, rocky outcroppings, etc.) or from an object sitting on the seabed. Three related measures of the level of speckle are commonly used in various literature: speckle contrast, scintillation index, and lacunarity. These are described and compared in Section 3.3.1, Section 3.3.2, and Section 3.3.3 respectively. Some applications of speckle analysis are then presented in Section 3.3.4.

3.3.1 Speckle contrast

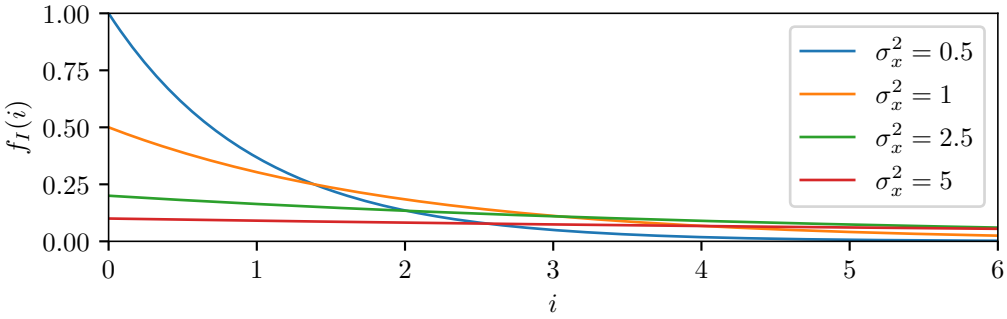
The speckle contrast is defined as the ratio of the standard deviation and the mean of the intensity [Goodman 1976],

$$C = \frac{\mathrm{Std}[I]}{\mathrm{E}[I]}. \tag{3.6}$$

From (3.4) the speckle contrast is unity for fully developed speckle. In other words, if the speckle contrast for a given region is one then that region meets the assumptions for fully developed speckle while other values indicate a departure from these assumptions.



(a) Speckle magnitude.



(b) Speckle intensity.

Figure 3.2: The probability density functions of (a) the magnitude and (b) the intensity of fully developed speckle for a range of variances. As the variance (proportional to the reflectivity of the seafloor) increases, the distributions become broader.

3.3.2 Scintillation index

Scintillation[2] is a generic term for a variation in the observed intensity of an object. It is commonly used in astronomy to describe the twinkling of stars due to the fluctuating refraction of light in a turbulent atmosphere [Osborn et al. 2015]. In physics it refers to a flash of light caused by a particle passing through a transparent material [Leo 1994]. In a radar sense, scintillation is used to describe the fluctuation in amplitude of a target from pulse to pulse [Skolnik 1990].

In sonar imagery, scintillation refers to the variations in the backscattered intensity of the scene. The commonly used measure of this variation is the scintillation index (SI), given by [Jackson and Richardson 2007]

$$\begin{aligned} \mathrm{SI} &= \frac{\mathrm{Var}[I]}{\mathrm{E}[I]^2}, \\ &= C^2. \end{aligned} \tag{3.7}$$

As this is the square of the speckle contrast, the same conditions apply: if it is unity, the image contains fully developed speckle, and other values indicate a deviation from the corresponding assumptions.

3.3.3 Lacunarity

The concept of lacunarity[3] arose in the field of fractal analysis [Mandelbrot 1982] where it was used as a measure of the spatial structure of binary-valued data. It has since been extended to analysis of multi-value data, with the lacunarity of a dataset given by [Plotnick et al. 1996]

$$\begin{aligned} L &= \frac{\mathrm{Var}[I]}{\mathrm{E}[I]^2}, \\ &= C^2, \\ &= \mathrm{SI}. \end{aligned} \tag{3.8}$$

As previously, a non-unity value for the lacunarity indicates the image is not fully developed speckle.

3.3.4 Applications

The preceding sections show that the speckle contrast, scintillation index, and lacunarity are effectively the same measure, differing only by a square root. The remainder of this thesis will use the lacunarity with the understanding that any comments can be equally applied to the other two.

[2] From the Latin *scintillāre*, to give off sparks or to glitter.
[3] From the Latin *lacūna*, meaning a gap or void.

As defined in (3.8), the lacunarity is an ensemble statistic, that is, a value that is calculated from all possible realisations of a scene. If the process is assumed to be ergodic (the ensemble averages are equal to the corresponding spatial averages [Birkhoff 1931]), the lacunarity can be calculated over a spatial window surrounding the point of interest. For fully developed speckle, the mean squared and the variance will be equal, meaning the calculated lacunarity $L = 1$. In the case of a constant value across the window, the variance will be zero and therefore the lacunarity will also be zero. Other values of L indicate a departure from the assumptions required for fully developed speckle. Note that this does not mean a given pixel is not speckle, but rather the collection of pixels in the window used to estimate the lacunarity do not meet the requirements.

In fractal analysis, the lacunarity is calculated for a variety of window sizes. Plotting it against the size of the windows shows the scales of repetition in a fractal image [Plotnick et al. 1996]. A similar approach with SAS data has shown that smooth and rippled seafloors exhibit different relationships as the window size changes, and that lacunarity analysis can be used to detect objects within sand ripples [Nelson and Kingsbury 2012; Nelson and Krylov 2014]. In the remote sensing community, lacunarity has been used in a variety of applications such as classifying land use from SAR images [Dekker 2003], feature selection from SAR imagery [Solberg and Jain 1997], and to classify different species of mangroves [Myint et al. 2008]. Williams [2015] has demonstrated the use of lacunarity to classify the nature of the seabed from SAS images allowing target detection performance to be predicted.

3.4 COHERENCE

The coherence is a complex measure of the similarity of two signals. For two complex signals s_1 and s_2 it is defined as [Born and Wolf 1999]

$$\Lambda = \frac{\mathrm{E}[s_1 s_2^*]}{\sqrt{\mathrm{E}\left[|s_1|^2\right]\mathrm{E}\left[|s_2|^2\right]}}, \tag{3.9}$$

where $\mathrm{E}[\cdot]$ is the expectation operator and * denotes the complex conjugate. The magnitude of the coherence $\delta = |\Lambda|$ is known as the degree of coherence and $\phi_0 = \arg\Lambda$ is the effective phase difference between the two signals. For two identical signals, i.e., $s_1 = s_2$, both the numerator and denominator of (3.9) collapse to $\mathrm{E}\left[|s_1|^2\right]$, and so $\Lambda = 1$ for identical signals. If the two signals are independent, then by definition $\mathrm{E}[s_1 s_2^*] = 0$ and so Λ is also zero for independent signals. Therefore, the degree of coherence is a value between zero and one indicating the similarity of two signals.

In other areas, the term coherence is used to refer to a spectral quantity [Shiavi 2010]. The definition is similar to (3.9) except that it is a function of frequency

rather than a single quantity from the time (or spatial) domain [Carter 1987]. This gives a measure of the similarity between different frequency components of the two signals [Manolakis, Ingle, and Kogon 2005].

3.5 SPATIAL ESTIMATION OF COHERENCE

Calculating the coherence requires knowledge of the ensemble statistics. This is not feasible for sonar imagery as it would require a large number of passes on identical tracks over different realisations of the scene. Instead, it is assumed that the data is ergodic, i.e., that the spatial average of a number of scatterers is equal to the ensemble average of a single scatterer [Birkhoff 1931]. This allows the expectations in (3.9) to be replaced with spatial averages, yielding the complex coherence estimator

$$\gamma = \frac{\sum_i s_1[i] s_2^*[i]}{\sqrt{\sum_i |s_1[i]|^2 \ \sum_i |s_2[i]|^2}}, \tag{3.10}$$

where i is used as an index over a finite two-dimensional region or window of M total pixels surrounding the point of interest. The magnitude of the estimate is the estimated degree of coherence, $d = |\gamma|$, and the argument of γ gives the interferometric phase ϕ between the images at that point. For the assumption of ergodicity to hold, the window must contain a number of independent scatterers with identical statistical properties. Note that there will not necessarily be M independent scatterers; if the image is oversampled (that is, the image resolution is higher than the system resolution) then the number of independent scatterers in the region N will be lower than the number of pixels M. Filtering applied during data processing and image formation can also increase the statistical dependency between adjacent pixels, thus further reducing N. The value of N is sometimes referred to as the effective number of looks (ENL). Although it can be estimated from system parameters, accurate values for a given dataset can be obtained via numerical methods such as maximum-likelihood estimation or the method of moments [Gierull and Sikaneta 2002].

The geometry of sonar imaging introduces a phase variation across the window. As illustrated in Figure 3.3, each point in the window has a different range to the sonar[4]. This corresponds to a phase offset in the imagery due to the different path lengths travelled by the echoes. To compensate for this, a phase correction term $\Delta\phi[i]$ is applied to the complex coherence estimator, giving

$$\gamma = \frac{\sum_i s_1[i] s_2^*[i] e^{j\Delta\phi[i]}}{\sqrt{\sum_i |s_1[i]|^2 \ \sum_i |s_2[i]|^2}}. \tag{3.11}$$

[4]Due to the way the coordinate system is chosen, this typically only varies in the across-track direction.

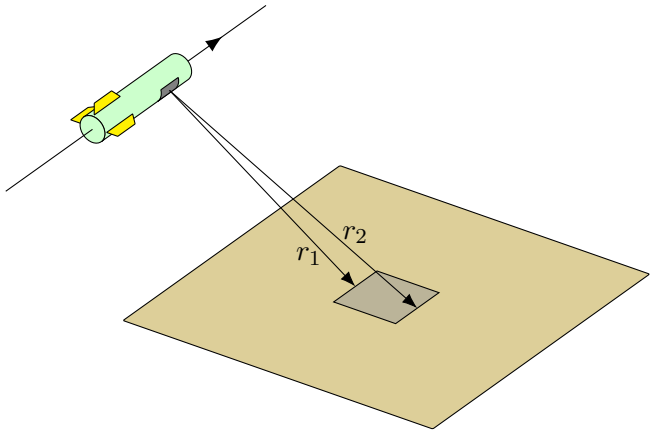


Figure 3.3: When estimating the coherence over the shaded spatial region, the varying path lengths across the region result in phase offsets in the imagery which need to be compensated.

Note that this phase correction term requires prior knowledge of the topography of the scene. For change detection applications, this may be available; for bathymetric or other applications it is unknown until later in the process. For this scenario, amplitude-only estimators such as the "quick and dirty" estimator [Guarnieri and Prati 1997] or a modified version based on the Siegert relationship [Touzi, Lopes, et al. 1999] are available to estimate the degree of coherence d without the need for phase correction. Throughout this thesis it is assumed that this correction has been applied to the imagery.

3.5.1 Coherence magnitude statistics

As sonar images are created by random processes, the estimated degree of coherence d must also be described statistically. Each measurement of d is an observation of the underlying random variable D. For speckle images with circular Gaussian statistics, the PDF of the distribution is [Touzi and Lopes 1996]

$$f_D(d;\delta) = 2(N-1)(1-\delta^2)^N d(1-d^2)^{N-2} \times {}_2F_1\big(N,N;1;\delta^2 d^2\big) \qquad 0 \le d \le 1, \tag{3.12}$$

where ${}_pF_q$ is the generalised hypergeometric function[5] and N is the number of independent scatterers in the estimation window. This PDF is plotted in Figure 3.4 for various values of the degree of coherence δ. This shows that d is biased towards higher values, with the bias being greater for smaller δ, i.e., the more dissimilar the two images are, the harder it is to predict the coherence. Figure 3.5 plots the PDF for $\delta = 0.95$ and a range of values for N. As might be expected, using more independent scatterers in the

[5]Gauss's hypergeometric function ${}_2F_1$ was the first hypergeometric function studied and is often referred to as 'the' hypergeometric function.

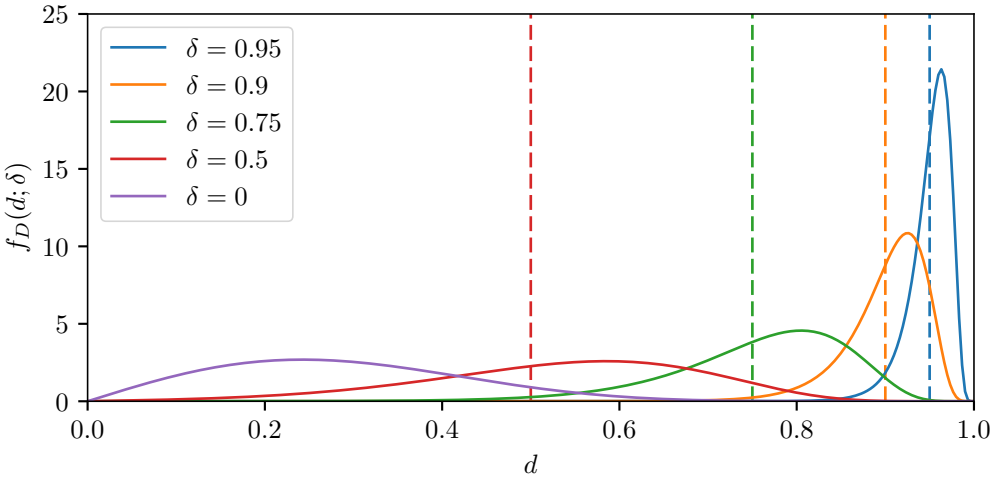


Figure 3.4: The PDF of the estimated degree of coherence d for various values of the true degree of coherence δ with a spatial window containing $N = 10$ independent scatterers. The vertical dashed lines mark the corresponding values of δ to highlight the bias in the estimate.

estimate results in a reduction in the bias and spread of the estimate.

The k-th moment of the distribution is given by [Touzi, Lopes, et al. 1999]:

$$\mathrm{E}\Big[D(d;\delta)^k\Big] = \frac{(1-\delta^2)^N \Gamma(N)\Gamma(1+k/2)}{\Gamma(N+k/2)} {}_3F_2\big(1+k/2, N, N; N+k/2, 1; \delta^2\big), \quad (3.13)$$

where Γ is the gamma function. The mean of the estimate $\mathrm{E}[D(d;\delta)]$ is plotted in Figure 3.6a for various values of N. This also shows the bias in the estimate, and its reduction with higher values of δ and N. The variance $\mathrm{Var}[D(d;\delta)] = \mathrm{E}\big[D(d;\delta)^2\big] - \mathrm{E}[D(d;\delta)]^2$ is shown in Figure 3.6b. As the degree of coherence and number of independent scatters used in the estimate increase, the variance gets smaller, i.e., the peak of the distribution gets sharper (for $\delta = 1$, it would be a Dirac delta).

Seymour and Cumming [1994] give the Cramér-Rao lower bound (CRLB) of the variance of the estimated degree of coherence as

$$\frac{\big(1-\delta^2\big)^2}{2N} \leq \mathrm{Var}[D(d;\delta)]. \quad (3.14)$$

This is plotted in Figure 3.7 for two values of N. This shows that the bound is accurate when N and δ are large, i.e., when the estimate is (relatively) unbiased, but as the bias increases the variance drops below the CRLB. As such, the CRLB should only be used if N and δ are large enough to minimise the bias, or if the bias in the estimate has been calculated and removed [Touzi, Lopes, et al. 1999].

Lyons and Brown [2013] considered the case where $\delta = 0$, i.e., the two images

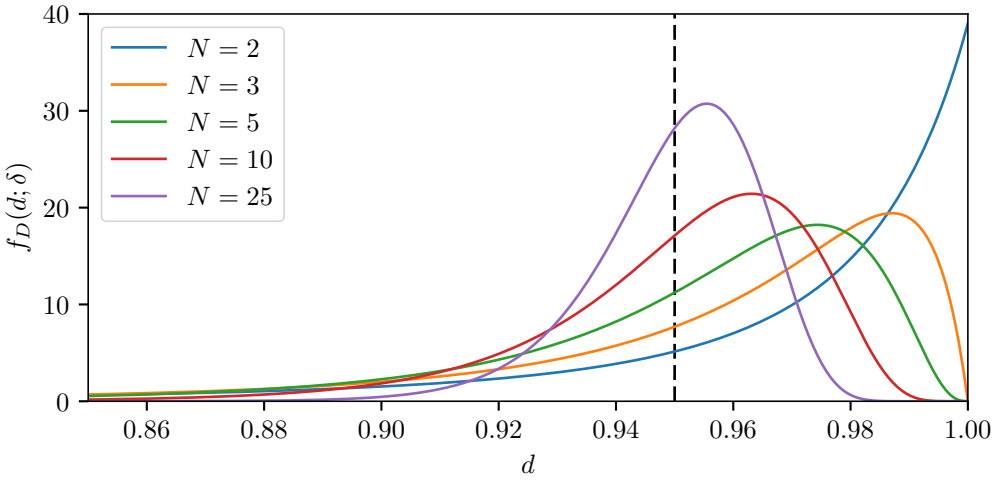


Figure 3.5: The PDF of the estimated degree of coherence d for $\delta = 0.95$ and increasing values of N (the number of independent scatterers in the estimation window). The vertical dashed line marks the value of δ.

contain uncorrelated speckle. They give the expected value of the estimated degree of coherence as

$$\mathrm{E}[D] = \frac{1}{N}\,\mathrm{E}\left[\left|\sum_{n=1}^{N} \frac{s_{n1}s_{n2}^*}{\sigma_{1n}\sigma_{2n}}\right|\right], \tag{3.15}$$

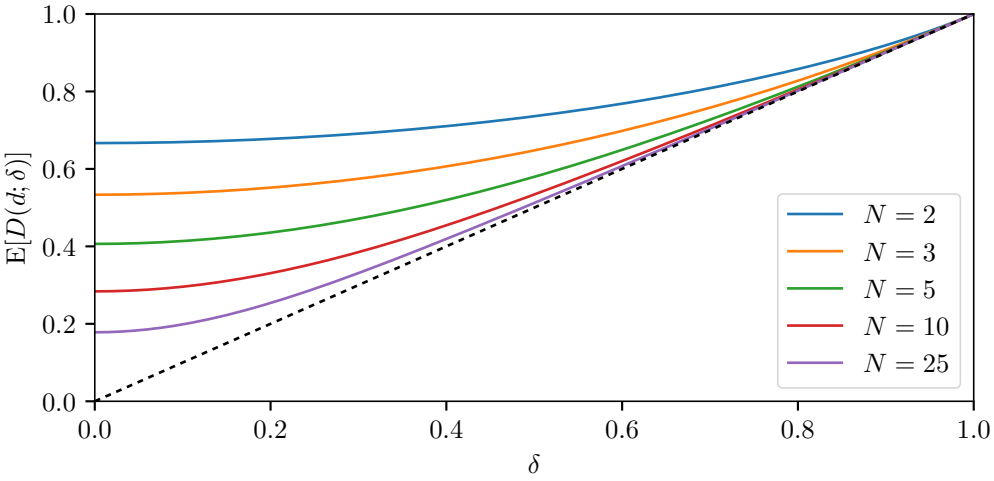
where σ_1 and σ_2 are the standard deviations of the images. Using the moment-generating function for the numerator given by Craig [1936], the expected value in the right-hand side of (3.15) is $\sqrt{N}$, and therefore $\mathrm{E}[D] = 1/\sqrt{N}$. Figure 3.8 plots this value against the corresponding expected value from (3.13). The $1/\sqrt{N}$ estimate is always higher than the value from the full model but provides a good approximation with a greater computational efficiency, albeit with a limited use-case.

3.5.2 Coherence phase statistics

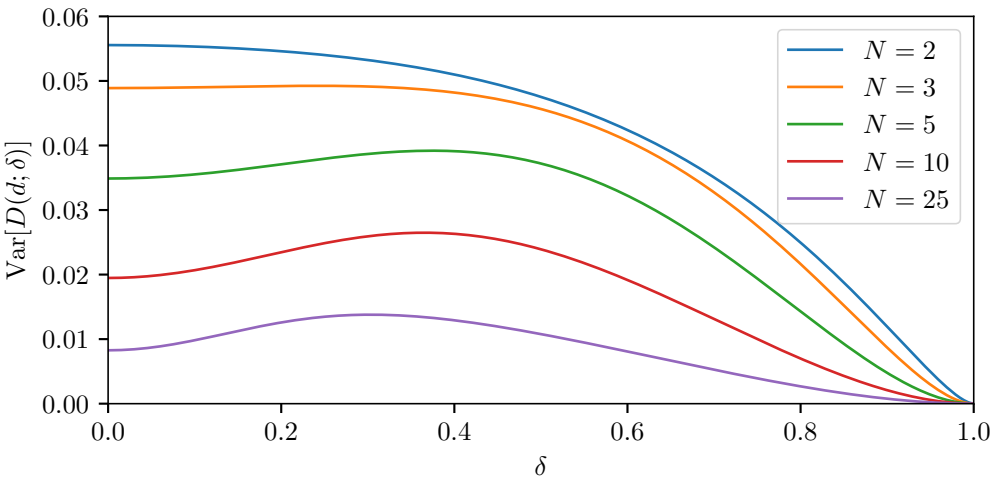
The estimated coherence phase is a random variable Φ. Goodman [1963] showed that the covariance matrix of interferometric data has a complex Wishart distribution, and using this Lee et al. [1994] derived the PDF of the estimated phase:

$$f_\Phi(\phi;\phi_0,\delta) = \frac{(1-\delta^2)^N}{2\sqrt{\pi}}\left(\frac{\Gamma(N+1/2)\beta}{\Gamma(N)(1-\beta^2)^{N+1/2}} + \frac{{}_2F_1\big(N,1;1/2;\beta^2\big)}{\sqrt{\pi}}\right) \qquad -\pi < \phi \leq \pi, \tag{3.16}$$

where $\beta = \delta\cos(\phi-\phi_0)$ and $\phi_0 = \arg\Lambda$ is the effective phase difference between the images as given by the ensemble coherence (3.9). In Figure 3.9, this PDF is plotted for



(a) Mean. The dotted line shows the true value of the degree of coherence.



(b) Variance.

Figure 3.6: The mean and variance of the estimated degree of coherence for various values of N.

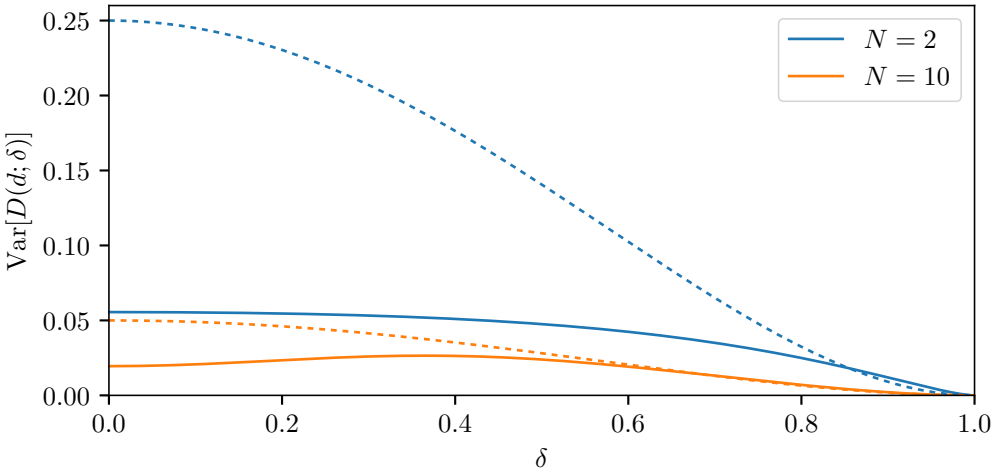


Figure 3.7: The variance of the estimated degree of coherence (solid lines) and the Cramér-Rao lower bound given by (3.14). For lower values of δ and N, i.e., where there is higher bias, the bound is inaccurate.

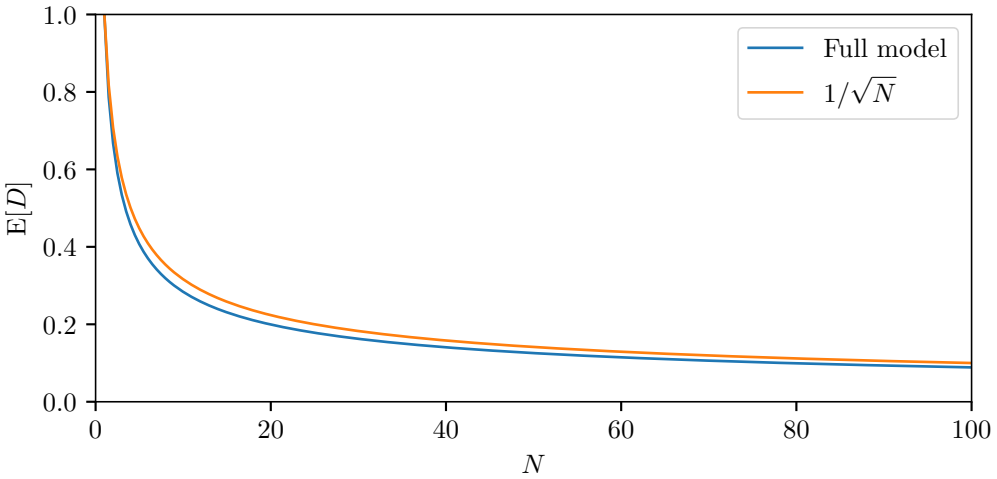


Figure 3.8: Comparison of the $1/\sqrt{N}$ approximation and the full model of (3.12) for the expected value of the estimated degree of coherence of uncorrelated ($\delta = 0$) speckle.

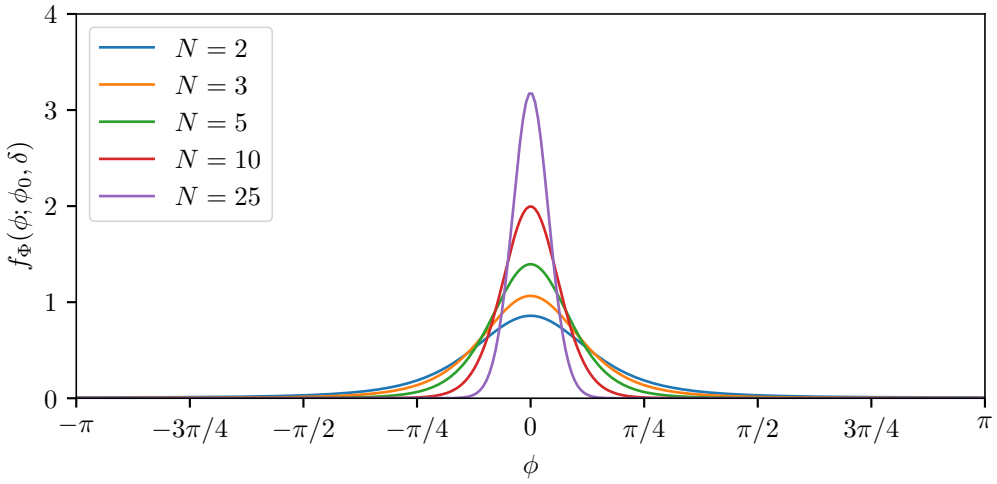


Figure 3.9: The PDF of the estimated coherence phase for various values of N. The degree of coherence $\delta = 0.75$ and the coherence phase $\phi_0 = 0$.

$\delta = 0.75$ and a number of values for N, while Figure 3.10 shows the effect of changing δ for a fixed number of scatterers $N = 10$. Unlike the magnitude estimate, the phase estimate is unbiased. As the degree of coherence and number of scatterers increase the distribution becomes tighter, approaching a Dirac delta as δ nears unity. For uncorrelated images ($\delta = 0$) the phase collapses to a uniform distribution. Figure 3.11 plots the variance of the phase estimator against the degree of coherence for a variety of values for N. As might be expected, the variance increases as δ and N decreases. In the $\delta = 0$ case the phase collapses to a uniform distribution on $(-\pi, \pi]$ which has a variance of $\pi^2/3 \approx 3.29$. All the plotted lines converge to this value as δ drops to zero. The CRLB of the phase estimate is given by Rodriguez and Martin [1992] as

$$\frac{1 - \delta^2}{2N\delta^2} \leq \mathrm{Var}[\Phi(\phi; \phi_0, \delta)]. \tag{3.17}$$

Figure 3.12 shows the variance for $N = 2$ and $N = 10$ plotted with the corresponding CRLBs. The CRLB is valid for high values of δ and N but becomes inaccurate as they decrease. This corresponds to the areas where the estimated degree of coherence d is biased. In the uncorrelated $\delta = 0$ case the CRLB tends to infinity.

For small values of N, mathematical identities can be used to replace the hypergeometric function with a combination of algebraic and trigonometric functions. For the $N = 1$ case, (3.16) becomes [Lee et al. 1994]

$$f_\Phi(\phi; \phi_0, \delta, N = 1) = \frac{1 - \delta^2}{2\pi} \frac{\sqrt{(1 - \beta^2)} + \beta(\pi - \arccos\beta)}{(1 - \beta^2)^{3/2}}. \tag{3.18}$$

By considering a simple interferometric system model, Just and Bamler [1994] inde-

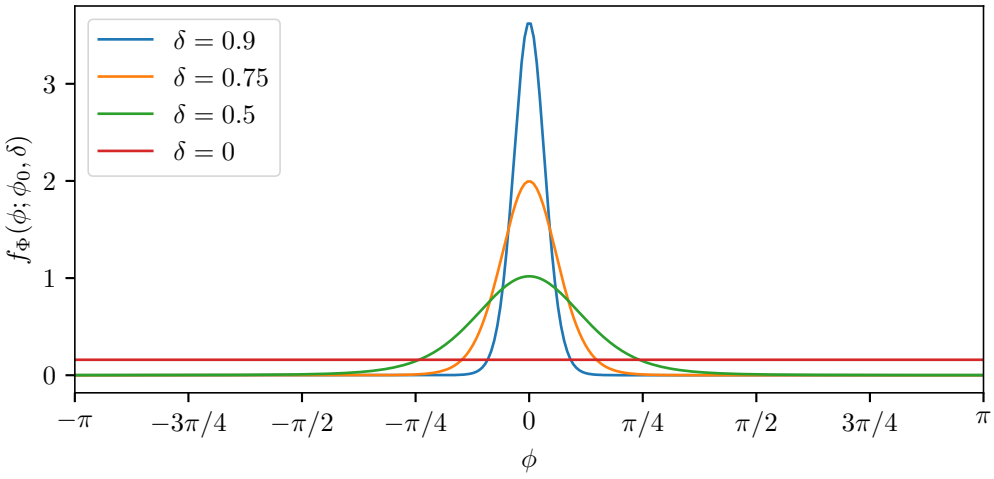


Figure 3.10: The PDF of the estimated coherence phase for various values of δ. The number of independent scatterers in the estimation region $N = 10$ and the coherence phase $\phi_0 = 0$.

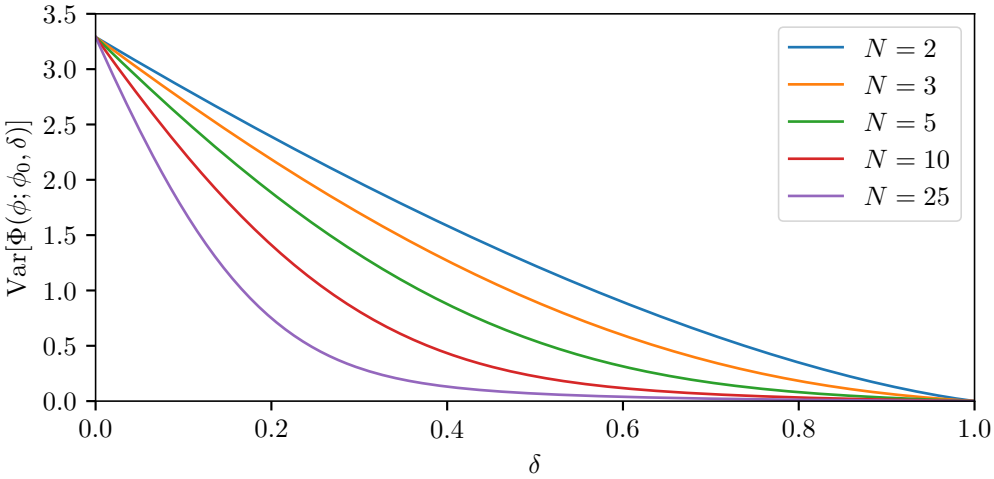


Figure 3.11: The variance of the estimated coherence phase for various values of N.

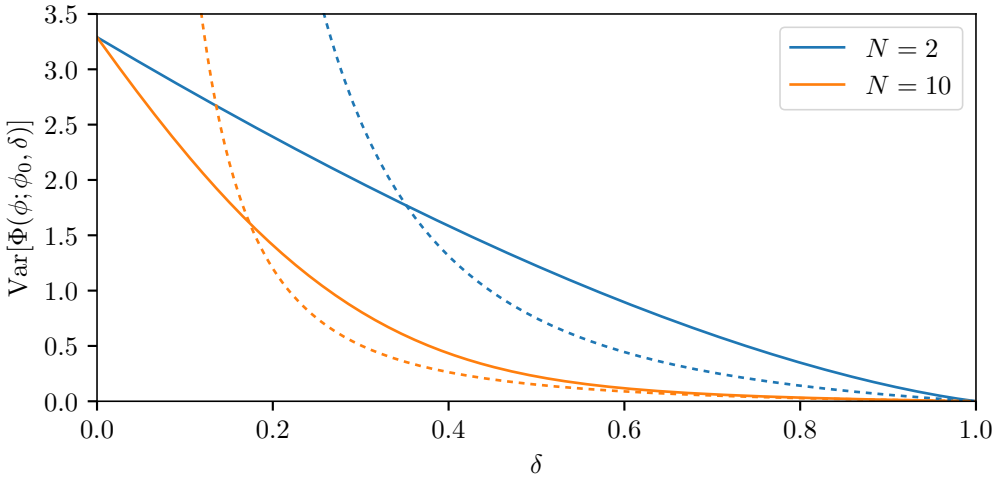


Figure 3.12: The variance of the estimated coherence phase (solid lines) vs the Cramér-Rao lower bound of (3.17) (dashed lines) for two values of N. The bound becomes inaccurate as the bias in the estimate increases.

pendently derived a PDF for the $N = 1$ case; with a suitable change of variables, this is equal to (3.18).

3.6 COHERENCE FACTORS

The coherence of a repeat-pass SAS system can be expressed as a product of factors [Bellec et al. 2005]

$$\gamma = \gamma_n \gamma_s \gamma_b \gamma_p \gamma_t, \tag{3.19}$$

where γ_n is the coherence factor associated with additive acoustic noise, γ_s is due to the footprint shift or image misregistration, γ_b is the result of baseline decorrelation, γ_p is caused by processing noise, and γ_t comes from the temporal decorrelation of the scene. In a ideal situation for change detection, γ_t will be one in regions where the scene did not change between passes, and zero where the scene differs. If the other coherence factors can be kept close to unity, then a loss of coherence directly corresponds to a change in the scene [Midtgaard et al. 2011]. The following sections give an overview of how each of the factors occurs and how it can be mitigated. For a more thorough treatment of these factors, the reader is directed to Barclay [2006].

3.6.1 Noise

The loss of coherence due to noise is governed by the SNR of the noise [Zebker and Villasenor 1992]. Assuming the noise is additive, the measured signals can be written

as

$$\begin{aligned} s_1 &= s + n_1, \\ s_2 &= s + n_2, \end{aligned} \tag{3.20}$$

where s is the common backscattered signal, and n_1 and n_2 is the noise in each measurement. Usually the noise is modelled with a Gaussian PDF. For sonar, the dominant noise source is the environment (i.e., acoustic noise) rather than electrical noise introduced at the receiver [Hayes and Gough 2008].

A significant contribution to this is noise from shipping, with propeller cavitation being a notable source at low frequencies [Ross 1976]. A study of commercial ships transiting the Santa Barbara channel showed an increase of up to 20 dB in acoustic noise between 10 Hz and 1 kHz at a range of 3 km, with the frequency and power increase dependent on both the type of ship and the heading of the ship relative to the receiver [McKenna, Ross, et al. 2012]. For some ships, the acoustic noise was detectable above the background levels at range of over 20 km. In open ocean areas like the Pacific basic, low frequency sounds can propagate for hundreds of kilometres [McKenna, Soldevilla, et al. 2009] while in coastal waters noise is dominated by local sources [McDonald, Hildebrand, Wiggins, and Ross 2008]. A quadrupling of the gross tonnage of commercial vessels between 1965 and 2003 has led the ambient noise due to ship activity increasing by over 10 dB in some locations [Andrew et al. 2002; McDonald, Hildebrand, and Wiggins 2006]. Platform noise — interference generated by the vessel operating or towing the sonar — is another acoustical noise source. Flow noise, caused by the movement of water past the receiving elements, is common to all moving sonar systems, while a hull-mounted sonar is also susceptible to mechanical vibrations coupled through the body of the vessel [Li 2012].

A variety of natural sources contribute to underwater acoustic noise. Breaking waves can increase noise levels when operating near the surf zone; Wilson, Wolf, and Ingenito [1985] observed an increase of 10 dB at 9 km range with the noise being detectable up to 15 km. Rainfall on the ocean surface can increase noise levels up to 35 dB across a range of frequencies into the tens of kilohertz [Nystuen 1986], and hail and snow are also detectable [Scrimger et al. 1987]. By adding a horizontal component to the velocity of the raindrops, wind increases the frequencies of the corresponding noise [Nystuen and Farmer 1987]. A large variety of marine life contributes to the acoustic noise. Species such as bottlenose dolphins use echolocation to find prey, while others such as whales communicate over long distances at low frequencies. Perhaps the most prevalent is the snapping shrimp (sometimes known as pistol shrimp) which generates broad peaks of noise at frequencies up to 200 kHz [Au and Banks 1998]. This was heavily studied during World War 2 due to its impact on the performance of military sonars[6] [University of California, Division of War Research 1946]. A report by the National Research Council [2003] provides a good overview into natural sources of

acoustic underwater noise.

For interferometric SAS employing multiple receiver elements in an array, a large portion of the noise will be correlated as it is common to all elements due to the acoustic noise dominating the electrical noise. In a repeat-pass scenario, the noise will be uncorrelated; although the sources may be the same, the actual noise sequence (realisation) received on each pass will be different. This allows the coherence estimator from (3.10) to be applied to give the acoustic noise factor as

$$\gamma_n = \frac{S}{\sqrt{S+\sigma_{n_1}^2}\sqrt{S+\sigma_{n_2}^2}}, \tag{3.21}$$

where S is the signal power and $\sigma_{n_1}^2$ and $\sigma_{n_2}^2$ are the variances of the noise in each measurement. If the two noise measurements have an identical variance $\sigma_{n_1}^2 = \sigma_{n_2}^2 = N$, where N is the noise power, this can be simplified to

$$\gamma_n = \frac{S}{S+N}. \tag{3.22}$$

The SNR is defined as

$$\mathrm{SNR} = \frac{S}{N}, \tag{3.23}$$

allowing the coherence factor to be written as [Rodriguez and Martin 1992]

$$\gamma_n = \frac{\mathrm{SNR}}{1+\mathrm{SNR}}. \tag{3.24}$$

A plot of this coherence factor as a function of SNR is given in Figure 3.13. This shows that an SNR of 1 (0 dB, i.e., where the signal and noise components have the same power) corresponds to a coherence of 0.5, and that an SNR of 20 dB or greater is required to keep γ_n above 0.99. If the coherence between two signals is known, then (3.24) can be inverted to give an effective SNR:

$$\mathrm{SNR} = \frac{\gamma_n}{1-\gamma_n}. \tag{3.25}$$

3.6.2 Footprint shift

The pulse transmitted by a sonar has an effective pulse duration T. For a narrow-band sonar T would be equal to the transmitted pulse duration, while for a wide-band sonar

[6]The noise created by snapping shrimp was often loud enough to drown out the propeller noise of submarines. Maps were distributed to U.S. submarine commanders showing areas where large beds of snapping shrimp might be found. These could then be used to hide from Japanese ships who largely relied on acoustic listening equipment to detect submarines. For his role in this, Dr. Martin Johnson was later awarded the National Academy of Sciences' Agassiz medal [*Sent shrimps into war for U.S.! Honored*, Chicago Sunday Tribune, March 15 1959]. There is also some evidence of bottlenose dolphins changing the frequency of their echolocation clicks to prevent them being masked by the noise from snapping shrimp [Au 1993].

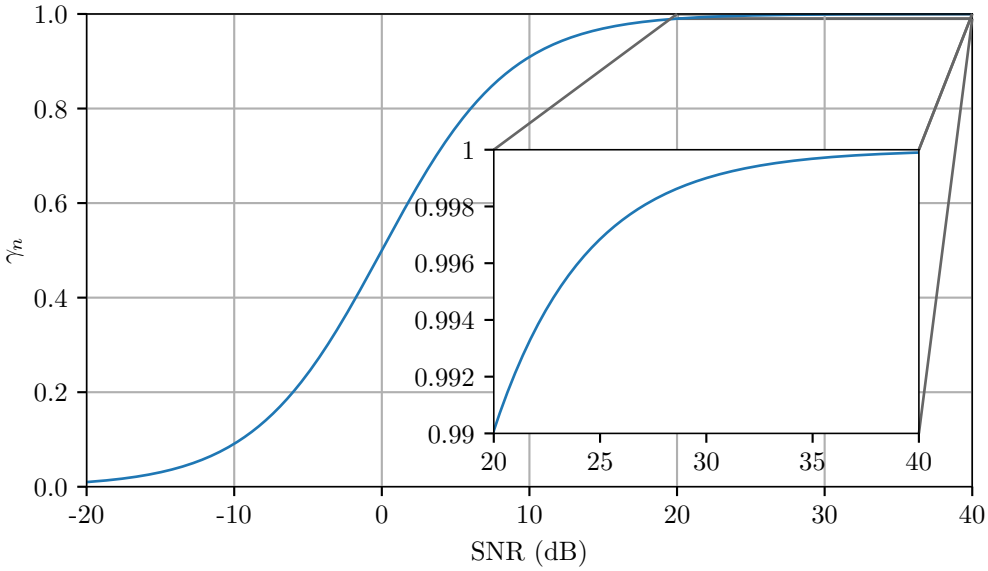
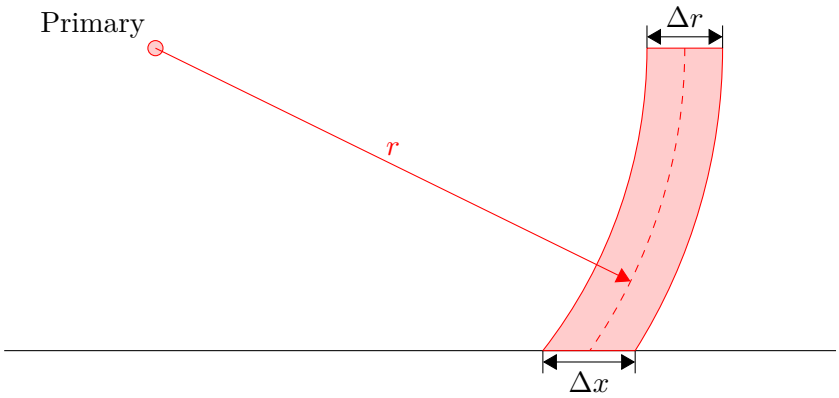


Figure 3.13: The coherence factor from the signal-to-noise ratio of additive white Gaussian acoustic noise.

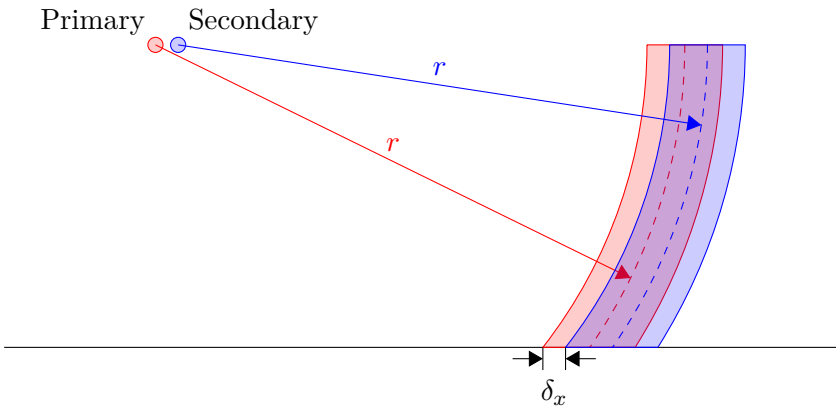
using pulse compression (as SAS does) the effective duration depends on the waveform and windowing applied, but can be estimated as $T \approx 1/B$ where B is the bandwidth of the modulated pulse. This corresponds to a given pulse width $\Delta r = cT$. As illustrated in Figure 3.14a, this means at a given range, r, the sonar is measuring the response of the scatterers within an annulus of width Δr. When projected onto the seafloor, this corresponds to a footprint of width Δx. An offset transducer, say from a second run over the scene, would result in the same range, r, corresponding to a different footprint location on the seafloor. As shown in Figure 3.14b, this results in a *footprint shift* of size δx. For an interferometric system using an array of hydrophones, the offset between the hydrophones can also cause a footprint shift. In both cases, the shift can be corrected by interpolating one of the responses to introduce a delay such that the footprints are aligned. This requires knowledge both of the relative position of the transducers and of the height above the seafloor. It also relies on the local tilt of the seafloor being known.

This footprint shift can be viewed as a misregistration of the two reconstructed images. If it is not corrected, then the corresponding pixels in each image will not be the result of the same spatial area in the scene. This in turn causes a loss of coherence; if the shift or misregistration is greater than the size of a sonar resolution cell then the coherence is completely lost. For a rectangular aperture, the coherence factor due to misregistration is given by [Just and Bamler 1994]

$$\gamma_s = \operatorname{sinc} \alpha, \tag{3.26}$$



(a) The footprint from a single pass.



(b) The footprint shift between passes.

Figure 3.14: The geometry of footprint shift. As shown in (a), the measured response at each range r corresponds to a footprint of width Δx on the seafloor; this width is a function of the effective duration of the transmitted pulse. On a second run (or when using multiple transducers on a single run), any offset in the measurement position results in the footprint shifting by δx. This is illustrated in (b). Accurate knowledge of the separation between passes allows the footprint shift to be corrected.

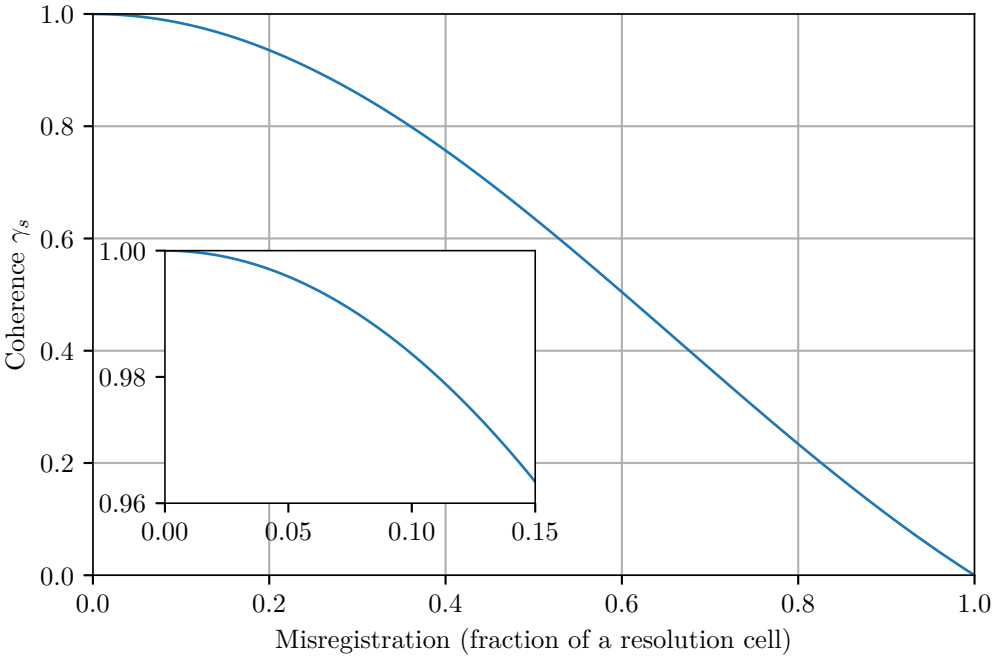


Figure 3.15: Coherence loss due to footprint shift or image misregistration.

where α is the misregistration as a fraction of a resolution cell size. This is plotted in Figure 3.15. A common consensus is that one-tenth of a resolution cell accuracy is required to minimise the loss of coherence from misregistration compared to other noise sources [Persons et al. 2002; Scheiber and Moreira 2000]. From (3.26), this corresponds to a coherence factor $\gamma_s = 0.984$, or an effective SNR of 17.8 dB.

3.6.3 Baseline decorrelation

Baseline decorrelation is also caused by the spatial separation between transducers; unlike the footprint shift, it can not be corrected [Zebker and Villasenor 1992]. As detailed in Section 3.1, the scattered signal in each resolution cell can be viewed as the superposition of multiple sub-resolution scatterers, each of which has a random amplitude and phase. When viewed from two separate positions, the speckle patterns of the backscattered signals differ, resulting in a loss of coherence [Jin and Tang 1996]. The degree of coherence loss depends on the *baseline* or distance between the receivers: the longer the baseline, the greater the baseline decorrelation. Another potential cause of baseline decorrelation is shown in Figure 3.16. Although the footprint shift can be corrected so the centre of the footprints are aligned, any difference in the height of two transducers means that the seafloor cuts the range annulus at different angles. As a result, the footprints are different sizes and the smaller will be imaging a subset of the scatterers that the larger footprint does. Therefore, the resulting interference patterns will be decorrelated.

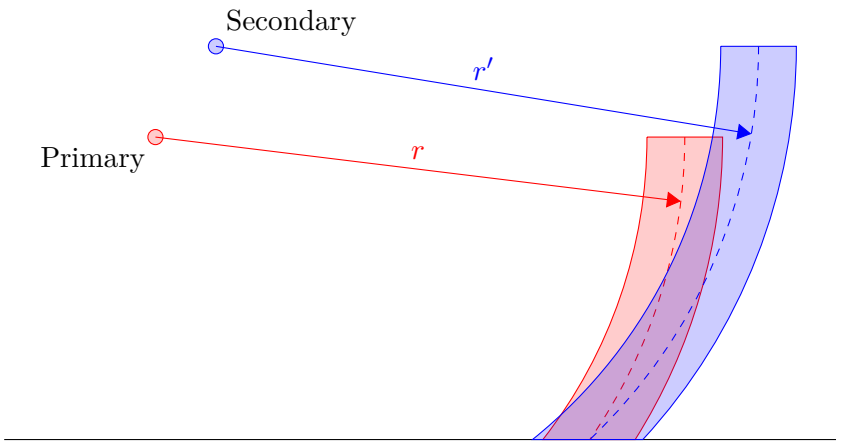


Figure 3.16: Baseline decorrelation due to different footprint sizes. The footprint shift has been corrected by interpolating the data from the secondary run so that the centres of the footprints are aligned. However, the different heights of the two runs means that the seafloor cuts the range annulus at a different angle, causing the footprints to be of different sizes.

The level of baseline decorrelation depends on the across-track shape of the resolution cell, $w(x)$, and the angles from the two hydrophones to the centre of the footprint, θ_1 and θ_2. It can be calculated as [Barclay 2006; Zebker and Villasenor 1992]

$$\gamma_b = \int_{-\infty}^{\infty} |w(x)|^2 \exp\left(-j\frac{4\pi f_0}{c}\ x\ \delta_\theta\ \cos\theta\right) \mathrm{d}x, \tag{3.27}$$

where f_0 is the centre frequency of the transmitted signal, $\theta = (\theta_1 + \theta_2)/2$ is the average angle to the centre of the footprint, and $\delta_\theta = \theta_1 - \theta_2$ is the difference in the angles. This can be seen as the Fourier transform of $|w(x)|^2$ evaluated for a spatial frequency $f_x = 2f_0/c\ \delta\theta\ \cos\theta$. For a system with a rectangular resolution cell the baseline decorrelation takes the form of a sinc function, and for a system with a sinc resolution cell it takes the form of a triangular function.

3.6.4 Processing noise

Reconstructing SAS images requires a number of processing steps, each of which can introduce noise into the result. Errors due to the limited precision of the storage format accumulate with each step in the process. The effect this has on the coherence is discussed in Section 3.6.4.1. Interpolation is another source of noise in the results and is covered in Section 3.6.4.2. For practical reasons, the synthetic aperture is commonly under-sampled. The artefacts introduced by this are quantified in Section 3.6.4.3.

3.6.4.1 Data precision

Sonar data is typically stored and manipulated using a floating point representation. As it is not possible to represent all possible real numbers, an error exists in (potentially) every stored number, and these errors accumulate as the data is processed [Goldberg 1991]. Each floating-point value v can be seen as representing an interval of numbers $[v-\epsilon, v+\epsilon]$, where ϵ is the precision of the floating-point representation being used. Applying interval arithmetic, adding two values v_1 and v_2 yields the interval $[v_1+v_2-2\epsilon, v_1+v_2+2\epsilon]$. For S additions, the worst-case error in the sum is proportional to S. For random data, the root-mean-square error grows with $\sqrt{S}$ [Higham 1993].

The IEEE-754 floating point representation [IEEE 2008] is commonly used for storage. It has a numerical precision

$$\epsilon_d = 2.2204 \times 10^{-16}. \tag{3.28}$$

After N_c calculations on a data point, the worst-case error will be $N_c\epsilon_d$, and the corresponding SNR is

$$\mathrm{SNR} \approx \frac{1}{N_c\epsilon_d}. \tag{3.29}$$

From (3.24) and Figure 3.13, an SNR of 40 dB gives a coherence factor $\gamma = 0.9999$. Applying this to (3.29) shows that $N_c \approx 4.5 \times 10^{13}$ calculations can be performed without any significant loss of coherence. As this is the number of calculations per data point, as opposed to over the entire dataset, the effect of data precision errors can be ignored for pratically all SAS processing [Barclay 2006].

3.6.4.2 Interpolation

Processing SAS data requires at least one interpolation step, and choosing an optimal interpolation method is critically important for minimising introduced noise. This noise can be analysed by considering the spectrums of both the signal and the interpolation kernel [Parker, Kenyon, and Troxel 1983]. Adapting this technique for synthetic aperture applications leads to a coherence factor [Hanssen and Bamler 1999]

$$\gamma = \frac{1}{\sqrt{1+\mathrm{SNR}^{-1}}} \times \frac{\int |H(f)|^2 I(f)\,\mathrm{d}f}{\sqrt{\int |H(f)|^2\,\mathrm{d}f \quad \int |H(f)|^2 |I(f)|^2\,\mathrm{d}f}}, \tag{3.30}$$

where SNR is the signal-to-noise ratio of the signal prior to the interpolation, $H(f)$ is the power spectral density of the signal, and $I(f)$ is the Fourier transform of the interpolation kernel. Barclay [2006] presents a study of a number of interpolation kernels, and shows that if the signal is oversampled by 2.5 times or greater then a linear interpolation introduces the least noise. This is because at these oversampling rates the passband distortion introduced by higher-order interpolators outweighs the aliased

noise. At lower oversampling rates, the higher-order interpolators provide a significantly better result. However, they are also computationally expensive. A common method to get around this is to initially oversample the signal with a fast interpolation (e.g., a sinc interpolation implemented by zero-padding in the Fourier domain), and then use an efficient linear interpolator to provide the final output.

3.6.4.3 Grating lobes

In a SAS system, the spatial sampling rate in the along-track direction is determined by the pulse repetition frequency and the forward velocity of the sonar. This finite sampling rate results in aliasing in the along-track spatial frequency domain. These aliased targets are incorrectly processed by the synthetic aperture reconstruction, leading to a phenomenon known as grating lobes [Hawkins 1996]. The ratio of the energy from the aliased targets to the energy from the real targets is known as the along-track ambiguity to signal ratio (AASR). This ratio can be estimated from the beampatterns of the transmitter and receiver, and the windowing applied as [Curlander and McDonough 1991]

$$\mathrm{AASR} \approx 10 \log_{10} \left(\frac{\displaystyle\sum_{\substack{m=-\infty \\ m\neq 0}}^{\infty} \int_{-\infty}^{\infty} \left[W\left(\frac{k_u}{B_p}\right) A(k_u + m k_{us}) \right]^2 \mathrm{d}k_u}{\displaystyle\int_{-\infty}^{\infty} \left[W\left(\frac{k_u}{B_p}\right) A(k_u) \right]^2 \mathrm{d}k_u} \right), \tag{3.31}$$

where $W(k)$ is the along-track windowing applied, $A(k)$ is the combined beampattern of the transmitter and receiver in the along-track direction, k_u is the along-track spatial frequency being evaluated, and k_{us} is the along-track spatial sampling frequency. The denominator is the total energy in the passband of the system, while the numerator is the sum of the energy that will be aliased into the passband due to the sampling. The coherence factor due to the AASR can be calculated using (3.24):

$$\gamma = \frac{1}{1 + \mathrm{AASR}}. \tag{3.32}$$

The along-track sampling rate is often specified in terms of the size of the transducer, i.e., $D/2$ sampling means the sample spacing is half the length of the transducer. Some early work claimed that a $D/2$ sampling rate resulted in the grating lobes being suppressed by a null in the along-track beampattern [Tomiyasu 1978]. However, this was subsequently shown to be based on incorrect analysis [Hawkins 1996] with (3.31) giving an AASR of -12.8 dB for $D/2$ sampling with rectangular apertures (i.e., sinc beampatterns) and no windowing.

The AASR and the corresponding coherence is shown in Figure 3.17. At low sampling rates the grating lobes are more pronounced, leading to a significant loss of

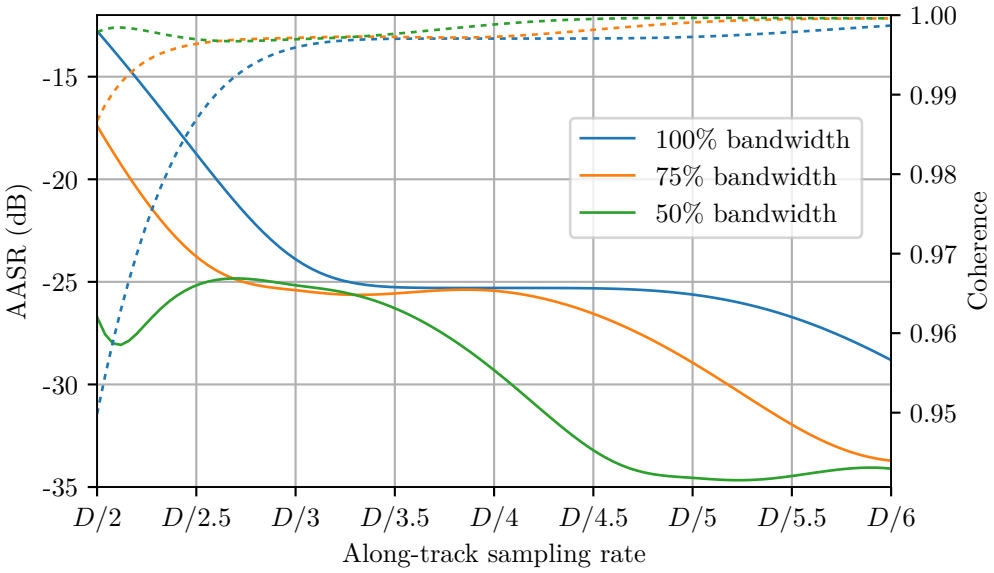


Figure 3.17: The AASR (solid lines) and the corresponding coherence (dotted lines) against the along-track sampling rate for a rectangular aperture. Reducing the bandwidth used in the reconstruction process results in a lower AASR due to reduced aliasing. However, this also results in a loss of the resolution of the reconstructed imagery.

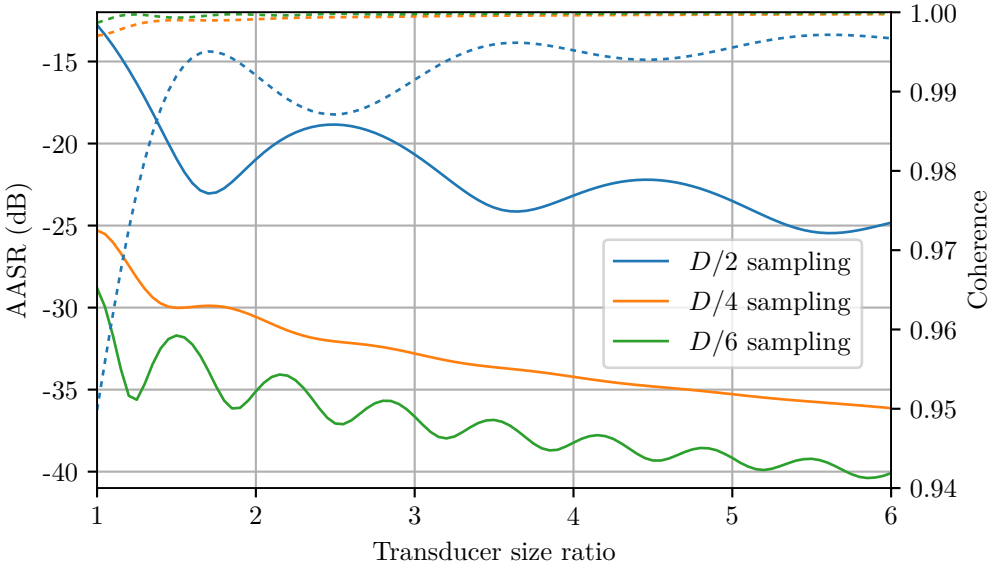


Figure 3.18: The AASR (solid lines) and the corresponding coherence (dotted lines) against the transducer size ratio for a rectangular aperture. Having one transducer be larger than the other means that nulls in their beampatterns are located at different positions thereby reducing the overall aliasing.

coherence. Using less of the available bandwidth in the reconstruction process (equivalent to making the window $W(k)$ narrower) reduces the level of aliasing and thereby the loss of coherence is also reduced. However, this also lowers the resolution of the reconstructed imagery. Another technique for reducing the AASR is to use different sized transducers for the transmitter and receiver. This results in the nulls in their beampatterns being located at different spatial frequencies which helps to minimise the aliasing. Figure 3.18 shows the AASR (and the resulting coherence) against the ratio of the transducer sizes for various along-track sample spacings.

3.6.5 Temporal decorrelation

Ultimately temporal decorrelation is what change detection attempts to measure. In an ideal scenario, the areas of a pair of images which have not changed will have a coherence of one, and areas that have changed will have a coherence of zero. However, if the coherence is low in all areas, then the result is ambiguous: either the entire scene has temporally decorrelated, or another source of decorrelation has corrupted the process e.g., the images are misaligned. To avoid this scenario, it is preferable to perform the repeat pass before the scene completely decorrelates. If there are distinct objects on the seafloor which are unlikely to significantly change (e.g., rocky outcroppings) then these act as marker points to check this alignment. However, for a bland seafloor this depends on the rate of change of the seafloor itself.

An obvious source of decorrelation is sediment transport. This is dependent on both the material and the location of the scene. In shallow coastal waters, for example, sediment is liable to be moved around by wave action. A study of two sites, one sandy and one silty, showed that the sandy scene decorrelated an order of magnitude quicker than the silty scene [Jackson, Williams, and Briggs 1996]. Imagery from a rail-mounted SAS system was used by Lyons and Brown [2013] to analyse the temporal change of a sandy seafloor in an area with a high amount of bottom-feeding fish. The presented results suggest that decorrelation occurs quicker for higher-frequency sonars, and for the shallow-water sandy sites studied the time until change detection is infeasible will be on the order of a few hours to a few days depending on the frequency of the sonar. An earlier model based around the diffusion equation [Jackson, Richardson, et al. 2009] was shown to provide a good estimate of the time taken to decorrelate.

In many areas fauna feeding and dwelling on the seafloor will cause decorrelation. One study observed changes on a sandy seafloor within 5 minutes, most likely caused by the numerous sand dollars (*Echinarachnius Parma*) at the site [Roderick, Dullea, and Syck 1984]. A change inside 13 minutes was noticed by Sæbø et al. [2011], while the change detection images presented by G-Michael, Marchand, Tucker, Marston, et al. [2016] show scattered areas of low coherence which would fit with this theory. Changes to the measured data can also be caused by other environmental effects such

as differences in properties of the water column [Hansen et al. 2015].

The acceptable time between passes will be dependent on the application, location, and environment. As shown by Midtgaard [2013], incoherent change detection is feasible over intervals longer than a year. On the other hand, the use of coherent techniques may require repeat passes to be performed within days.

3.6.6 Summary

The coherence factors discussed in this section have different causes and mitigation strategies. Baseline decorrelation (γ_b) is a direct result of the viewing geometry. The spatial separation between the path followed by the sonar on each is a function of the accuracy of the navigation system used to position it. The more accurate the positioning is, the closer the paths will be, and the less effect the baseline decorrelation will have. Acoustic noise (γ_n) is an unavoidable property of the medium[7] and cannot be subsequently improved. Electrical noise, although generally insignificant compared to the acoustic noise, can be minimised through good design of the receiver electronics. Temporal decorrelation δ_t cannot be mitigated, although for change detection applications this is less of an issue than for interferometry as measuring it is the aim. The limitation here is whether there is enough common information to allow the two images to be registered with sufficient accuracy to give meaningful results. The effects of footprint shift or image misregistration (δ_s) can be minimised, if not completely eliminated, by accurate processing of the data. Similarly, the processing noise δ_p is unavoidable but can again be minimised by careful implementation of the applied algorithms.

3.7 CORRELATION

In general usage, correlation describes a mutual relationship between two or more objects. From a signal processing viewpoint it is a measure of the linear dependence between two random signals as a function of the displacement between them [Clapham and Nicholson 2009]. Closely related to convolution, it is often used to search for common features between two signals or images. Correlation can also be defined for random variables. Although not directly used in this thesis, the statistical definitions are presented in Appendix C for completeness.

[7] Technically, the acoustic noise created by the vessel operating the sonar can be reduced, but this may result in other components worsening. For example, slowing the vessel down reduces the effectiveness of the control, and thus the error in its positioning increases.

3.7.1 Correlation of two signals

For two continuous, wide-sense stationary functions $x(t)$ and $y(t)$, the covariance is defined as [Orfanidis 1989]

$$\mathrm{cov}_{xy}(\tau) = \int_{-\infty}^{\infty} (x(t) - m_x)^* \, (y(t+\tau) - m_y) \, \mathrm{d}t, \tag{3.33}$$

where m_x and m_y are the means of $x(t)$ and $y(t)$ respectively, and τ is the displacement of the second signal relative to the first and is commonly referred to as the lag. For two discrete signals $x[t]$ and $y[t]$, the definition is analogous:

$$\mathrm{cov}_{xy}[\tau] = \sum_{n=-\infty}^{\infty} (x[n] - m_x)^* \, (y[n+\tau] - m_y), \tag{3.34}$$

where τ is the displacement as an integer number of samples. This is a measure of how similar the variations in the signals are. If both of them are positive at the same time, or both of them are negative at the same time, then the product will always be positive and the overall correlation will be large. If they have the opposite signs — one is negative when the other is positive — the product will always be negative and so will the overall correlation. If there is no such relationship, then the product will vary between positive and negative and the correlation will tend to zero. For complex signals, the usage of the conjugate ensures that aligned features in the imaginary components of the signal contribute positively to the covariance [Therrien 1999].

It is difficult to infer the level of correspondence between the signals solely from the covariance. If both signals are multiplied by some constant a, then from (3.33) the covariance is

$$\begin{aligned} \mathrm{cov}_{ax,ay}(\tau) &= \int_{-\infty}^{\infty} (ax(t) - am_x)^* \, (ay(t+\tau) - am_y) \, \mathrm{d}t, \\ &= a^2 \, \mathrm{cov}_{xy}(\tau). \end{aligned} \tag{3.35}$$

Although the relationship between the variables has not changed, the value of the covariance has. To avoid this uncertainty, the covariance can be normalised by the standard deviations of the signals to yield the correlation function:

$$\mathrm{corr}_{xy}(\tau) = \frac{1}{\sigma_x \sigma_y} \int_{-\infty}^{\infty} (x(t) - m_x)^* \, (y(t+\tau) - m_y) \, \mathrm{d}t \tag{3.36}$$

for continuous signals, and

$$\mathrm{corr}_{xy}[\tau] = \frac{1}{\sigma_x \sigma_y} \sum_{n=-\infty}^{\infty} (x[n] - m_x)^* \, (y[n+\tau] - m_y) \tag{3.37}$$

for discrete signals. This ensures that the magnitude of the correlation is a value in the closed interval $[-1, 1]$. A value of 1 indicates that the two signals have a perfect positive linear relationship (when one increases so does the other, and vice-versa). Similarly, value of -1 indicates a perfect negative linear relationship (when one increases the other decreases, and vice-versa). If the correlation is zero then the variables are said to be uncorrelated: there is no linear relationship between them.

It is worth noting that in some signal processing literature the correlation is defined as the unnormalised version, i.e., the same as the covariance. The autocorrelation is the correlation of a signal with itself, while the cross-correlation is the correlation between two separate signals. In this thesis, a reference to correlation implies a cross-correlation; again, other literature may differ in this regard.

3.7.2 Correlation of sampled signals

For a finite sampled signal, (3.37) must be adjusted to account for the limited data available. Suppose the first signal $x[n]$ has N samples, and the second signal $y[m]$ has M samples. The second signal can be extended to be zero outside the measured samples, i.e.,

$$y'[m] = \begin{cases} y[m] & 0 \leq m < M, \\ 0 & \text{otherwise.} \end{cases} \tag{3.38}$$

The correlation is then given by

$$\mathrm{corr}_{xy}[\tau] = \frac{1}{\sigma_x \sigma_y} \sum_{n=0}^{N-1} (x[n] - m_x)^* \, (y'[n+\tau] - m_y). \tag{3.39}$$

Typically, the signal standard deviations are not known but instead are calculated from the samples, yielding

$$\mathrm{corr}_{xy}[\tau] = \frac{\sum_{n=0}^{N-1} (x[n] - m_x)^* \, (y'[n+\tau] - m_y)}{\sqrt{\sum_{n=0}^{N-1} |x[n] - m_x|^2} \sqrt{\sum_{m=0}^{M-1} |y[m] - m_y|^2}}. \tag{3.40}$$

This is defined for any integer lag τ. However, for the majority of lags there will be no overlap of the signals and, due to the zero extension applied by (3.38), the correlation at these points will be zero. In total, there will be $N + M - 1$ values of τ for which there is some overlap between the non-zero portions of the signals. For some of these

lags, there will only be a partial overlap. This reduces the total energy that can appear in the numerator of (3.40). As the denominator is constant, the maximum value that the correlation can achieve is reduced in these edge regions.[8]

3.7.3 Relationship to coherence

For zero-mean signals — such as speckle images — the correlation (3.40) collapses to

$$\mathrm{corr}_{xy}[\tau] = \frac{\sum_{n=0}^{N-1} x[n]^* \, y'[n+\tau]}{\sqrt{\sum_{n=0}^{N-1} |x[n]|^2} \sqrt{\sum_{m=0}^{M-1} |y[m]|^2}}. \tag{3.41}$$

Comparing this to the coherence estimator of (3.10) shows that the coherence is equivalent to a zero-lag correlation. Alternatively, the correlation at a given lag corresponds to the coherence that would be estimated if the second signal were to be shifted by that amount. This means that the statistics of the coherence estimator can also be applied to the correlation. For example, from (3.12) the PDF of the correlation magnitude estimate at a lag τ would be $f_D(d; \delta_\tau)$ where δ_τ is the true correlation magnitude at that lag.

3.7.4 Limitations

It should be borne in mind that the correlation coefficient is only a measure of the linear relationship between the two variables. Non-linear relationships are not guaranteed to be detected by the correlation. For example, if $y(t) = x^2(t)$ it is obvious that $y(t)$ is wholly dependent on $x(t)$. However, the correlation is zero in this case.[9] It follows from this that a high correlation does not necessarily mean that the relationship between the signals is linear, but rather to what degree it can be approximated as linear.

3.8 EFFICIENT IMPLEMENTATION OF WINDOWED STATISTICS

The direct approach of calculating windowed statistics involves a sliding window technique: the window is placed on the image, the corresponding statistic calculated, and the window is then shifted by a pixel for the next calculation. For a $M \times N$ pixel image and a $P \times Q$ window, there are $(M - P + 1) \times (N - Q + 1)$ windows. Even for a simple statistic, the amount of computation required grows quickly with the image size.

[8] Signal processing software libraries often give the user a choice of what portion of the correlation to calculate. The region where there are complete overlaps between the signals is commonly referred to as the "valid" region, while the complete correlation including these edge effects is the "full" region.

[9] This is shown in Section C.3 for the statistical correlation.

For example, if the pixels in the window are being summed, then the inner operation has a big-O complexity $\mathcal{O}(PQ)$ and the overall calculation is $\mathcal{O}(MNPQ)$. Eliminating some of the nested loops in the evaluation of these statistics would result in a large performance improvement.

3.8.1 Integral images

Summed area tables were introduced to computer graphics by Crow [1984] and then to computer vision by Lewis [1995]. The name 'integral image' was coined by Viola and Jones [2004] to distinguish its use in image analysis as opposed to the original texture mapping. For an image X, the integral image II is defined as

$$II[r, c] = \sum_{r'=0}^{r} \sum_{c'=0}^{c} X[r', c'], \tag{3.42}$$

where r and c are the row and column indices respectively. In other words, the value of the integral image at a given pixel $[r, c]$ is the sum of the pixels above and to the left of the corresponding pixel in the source image (including the pixel itself, and all pixels in its row and column).

Having computed the integral image, it can then be used to efficiently calculate the sum of pixels in an arbitrary rectangular area of the source image. If the area of interest is defined by the top-left corner $[m, n]$ and the bottom-right corner $[p, q]$ (with the corners included in the area), the direct way of calculating the sum S is

$$S = \sum_{r=m}^{p} \sum_{c=n}^{q} X[r, c]. \tag{3.43}$$

Instead of this looping process, the integral image can be used to calculate the sum with four lookups and three additions:

$$S = II[p, q] - II[p, n-1] - II[m-1, q] + II[m-1, n-1]. \tag{3.44}$$

As illustrated in Figure 3.19, this can be viewed as taking the area of the entire image up to and including the bottom-right of the desired window, and then removing the areas above and left of the window. Note that (3.44) is independent of the size of the window, i.e., once generated, the integral image can be used to find the sum of any size region in the original image.

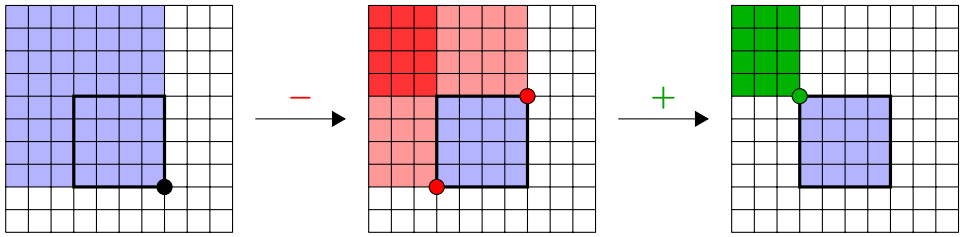


Figure 3.19: Calculating the sum of the marked spatial region from an integral image. The initial value is the value of the integral image at the bottom-right of the region. The areas to the left and top of the window are then removed by subtracting the integral image values at the bottom-left and top-right respectively. Since this subtraction results in the area above and left of the window being removed twice, the final step is to add that back in to obtain the sum.

3.8.2 Lacunarity

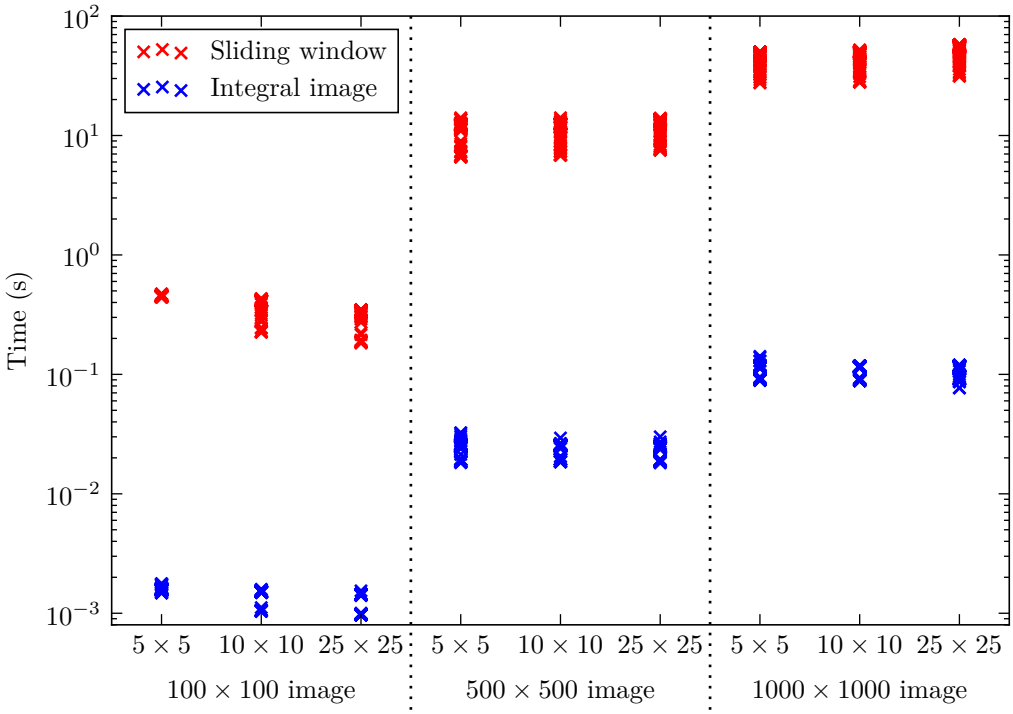
By using the definition $\mathrm{Var}[I] = \mathrm{E}\left[I^2\right] - (\mathrm{E}[I])^2$, the calculation of lacunarity as per (3.8) can be implemented as

$$\begin{aligned} L &= \frac{\mathrm{Var}[I]}{\mathrm{E}[I]}, \\ &= N \frac{\sum_i I[i]^2}{(\sum_i I[i])^2} - 1, \end{aligned} \tag{3.45}$$

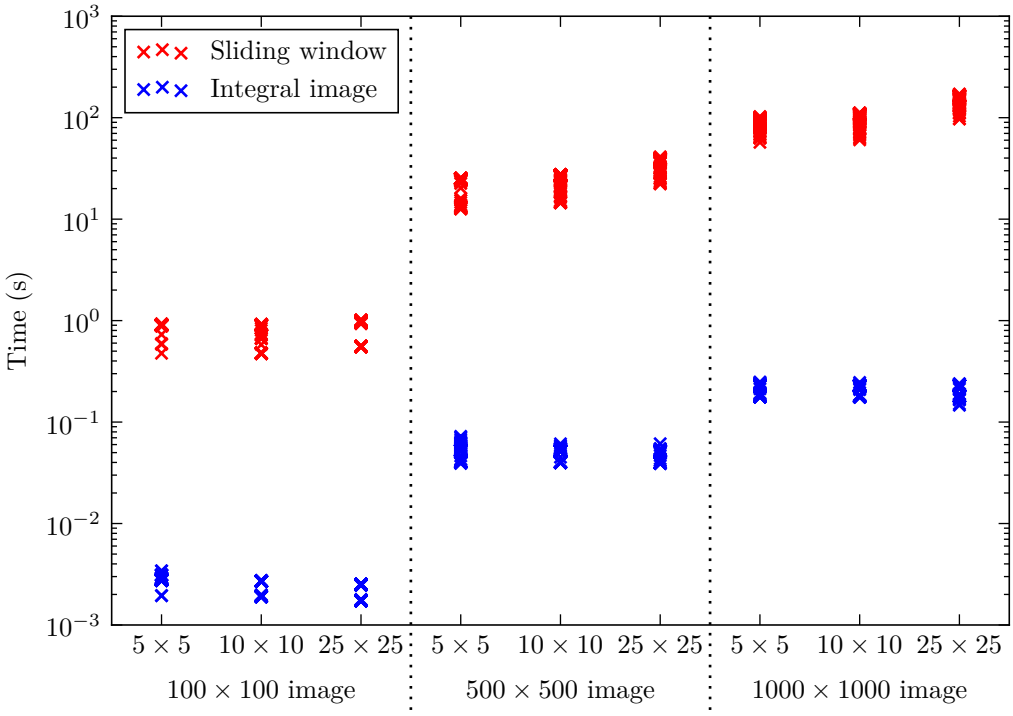
where i is used as an index over the window being evaluated. If this is performed with sliding windows, the inner calculation involves squaring all the values (PQ multiplications), two sets of sums ($2PQ$ additions), and the final squaring, division and subtraction (3 operations). This requires $3PQ + 3$ elementary operations for order $\mathcal{O}(PQ)$ complexity, corresponding to $\mathcal{O}(MNPQ)$ to perform it over the entire image.

As demonstrated by Williams [2015], integral images can be used to reduce the complexity of the calculation. This requires forming two integral images; one for the intensity and one for its square. The sums for each of these can then be efficiently computed using (3.44) for the desired window size, and then the lacunarity value can be calculated per (3.45). Between generating the integral images, computing the sums, and the final squaring, division and subtractions, this requires 10 elementary operations per pixel, and thus the complexity is $\mathcal{O}(MN)$.

Figure 3.20a shows the time required to calculate the lacunarity for a variety of image and window sizes. The median speed increases and maximum relative errors from these trials are given in Table 3.1. From this it can be seen that the speedup from using integral images is over two orders of magnitude with no appreciable loss of accuracy for the image sizes used.



(a) Lacunarity.



(b) Spatial coherence estimate.

Figure 3.20: The time taken to calculate (a) the lacunarity and (b) the spatial coherence estimate by the sliding window and integral image techniques. Each data set comprises 100 trials. The labels on the x-axis are the window sizes the statistic was calculated over. Note that the y-axes are plotted on a log scale.

		Median speedup		Maximum error	
Image	Window	Lacunarity	Coherence	Lacunarity	Coherence
100×100	5×5	287	321	2.3×10^{-12}	9.2×10^{-13}
	10×10	273	333	4.2×10^{-13}	9.4×10^{-13}
	25×25	235	385	6.7×10^{-14}	2.3×10^{-12}
500×500	5×5	483	474	1.1×10^{-10}	4.6×10^{-11}
	10×10	493	510	1.1×10^{-11}	2.2×10^{-11}
	25×25	541	740	2.3×10^{-12}	4.3×10^{-12}
1000×1000	5×5	409	409	3.9×10^{-10}	6.6×10^{-10}
	10×10	423	454	5.1×10^{-11}	7.7×10^{-11}
	25×25	470	680	8.8×10^{-12}	1.9×10^{-11}

Table 3.1: The median speedup and maximum error resulting from using integral images for calculating the lacunarity and spatial coherence estimate instead of the sliding window technique. The results are from 100 trials run in Python using the NumPy library. The error is the maximum difference in the calculated values in any window over all trials and is relative to the sliding window value.

3.8.3 Spatial coherence estimation

The estimation of the spatial coherence is another candidate for the use of integral images for performance improvements. Using sliding windows to implement the estimator (3.10) requires $3PQ+3$ elementary operations to calculate the estimated value for each spatial region, resulting in a computational complexity $\mathcal{O}(MNPQ)$ for the whole image. For the integral image case, three images are required: one for the conjugate product in the numerator, and two for the normalising factors in the denominator. Overall, $21MN$ operations are required for the calculation so the complexity is $\mathcal{O}(MN)$.

Calculation times for the spatial coherence estimate using both the sliding window and integral image techniques are shown in Figure 3.20b for a range of image and window sizes. The corresponding median speed increases and maximum relative errors are presented in Table 3.1. As with the lacunarity calculation, the use of integral images for estimating the spatial coherence results in a significant two-order-of-magnitude speed increase while not introducing any significant errors.

3.8.4 Floating point error accumulation

As outlined in Section 3.6.4.1, the floating point system typically used to store SAS data cannot represent every possible real number accurately, and thus errors are introduced. From Table 3.1 it can be seen that the maximum relative error[10] in the integral

[10]There will obviously also be floating point errors in the sliding window calculation. Since each window is summed independently, S never gets very large and so the errors are minimal. As such errors are inevitable with floating point applications, the sliding window values are considered the 'true' value with which to compare the integral image outputs.

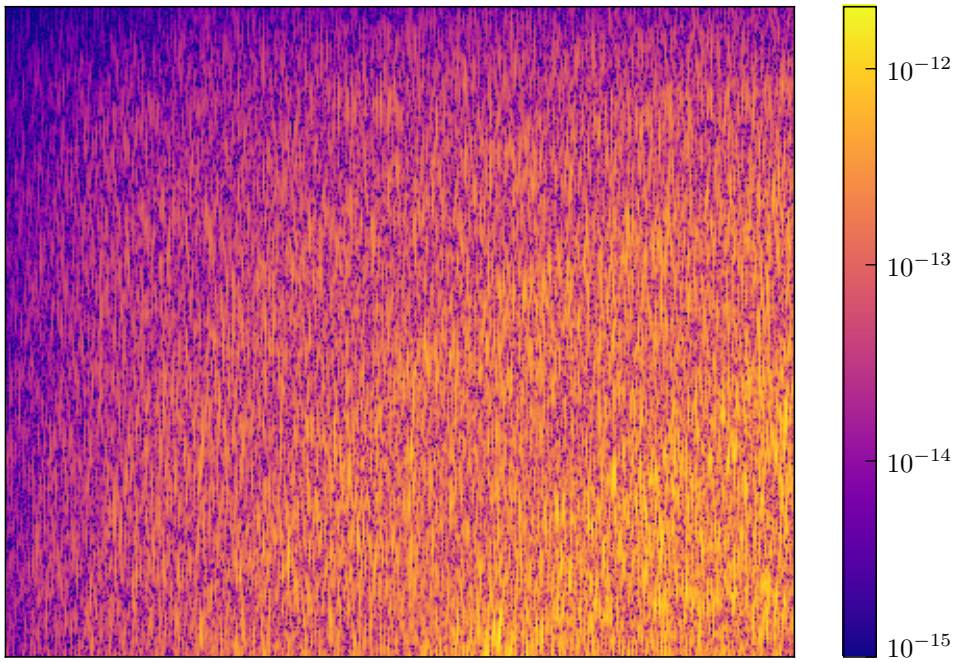


Figure 3.21: The error in the lacunarity due to use of an integral image in the calculation. The values are the relative error compared to the corresponding sliding window value. The image is 500×500 pixels and the lacunarity was calculated in 25×25 pixel windows. The integral image origin was the top-left pixel.

image method is bigger for larger images, i.e., as the number of pixels being accumulated in the integral image increases. Figure 3.21 shows the error over a 500×500 pixel lacunarity map. As would be expected, the maximum error occurs in the corner diagonally opposite the integral image origin as this is where the number of additions need to calculate the integral image value is at its maximum. The precision of the floating-point implementation also affects the maximum error; Williams [2015] reports double-precision data types having sufficient accuracy when using integral images to calculate the lacunarity from source images containing over 14 million pixels.

The error introduced by the use of integral images with a given system (determined by the resolution, image sizes, etc.) should be able to be characterised from a number of sample images. Alternatively, this can be estimated on the fly by also using the sliding window technique to calculate the final value and comparing to the integral image output (as the processing errors are randomly distributed, it may be necessary to perform this test over a few outputs to avoid outliers). If this error proves to be too large, then a straightforward solution is to sub-divide the image and process each piece independently. Although this involves overlapping the integral images to avoid gaps in the final output, the amount of extra computation is still negligible compared to the sliding window technique. An alternative method would be to use the centre of the image as the origin rather than a corner, and integrate outwards. While this would

halve the number of accumulations (and thus the error) needed to obtain the values at the edge, it would complicate the logic needed to recover the sum of a given block. The corner furtherest from the origin would need to be determined to use as the starting point for (3.44), and similar calculations would be needed for the other corners. Blocks which span the origin would have to be divided into four sub-blocks.

3.9 DISCUSSION

Speckle is a fact of life for SAS imaging. Its presence requires the reconstructed data to be treated as random variables, and thus any subsequent analysis of the data must also be undertaken statistically. The intensity and phase of the speckle follows well known, simple distributions, and this fact can be utilised for applications such as object detection and seafloor classification. The distribution of the coherence between two speckle images is also well known, albeit less simple. The dimensions of the spatial region used in the estimation of the coherence directly affects the distribution, and thus its choice is critical for optimal results. The correlation between two regions of speckle imagery follows the same distributions and so the same care must be taken in the choice of window size.

Many factors have effects on the coherence observed with repeat-pass imagery. From the point of view of change detection, most of these are unavoidable: once the data is collected, they cannot be compensated for. The main exception to this is the footprint shift or image misregistration factor. This can be improved by either correcting the navigation data and regenerated the images, or by directly warping the repeat-pass image to align with the primary pass. Coherent change detection techniques require a high level of accuracy in the registration of the two images meaning this step is vital.

Chapter 4

REPEAT-PASS SAS SIMULATION

Given the cost of performing sea trials, simulated data is widely used in the development of sonar processing algorithms. For repeat-pass applications the need for an accurate ground truth to check registration procedures against also requires simulation. One simulation technique is to generate the received images directly. This is typically done by starting with noise from an appropriate statistical distribution and applying a filter based upon the properties of the sonar system. Although this is a simple approach, it is limited in that it generally only works for noise (speckle) images and cannot generate results from a more complex scene. For this reason, most simulation approaches model the scene and generate the acoustic field that would have been observed at the receiver. This data can then be processed and reconstructed as it would for a real-world system.

One of the simplest models is the point scatterer model in which the scene is decomposed into a collection of infinitesimally small points. Although their lack of size implies they would not scatter any energy, each point is assigned a non-zero scattering strength on the basis it represents a portion of a larger scattering object. Any energy incident on the point is then scattered equally in all directions. The scattering strength may be arbitrarily chosen, or it can be calculated using the Rayleigh scattering theory [Rayleigh 1896] for a sphere modelling the portion of the object represented by the scatterer. Although simple to implement, point scattering is inefficient: on the order of ten scatterers per wavelength (corresponding to one hundred per square wavelength of area, or one thousand per cubic wavelength of volume) are required for adequate modelling of a continuous surface [Hunter 2006].

A similar technique is to describe the simulated scenes analytically; this allows a straightforward evaluation of the received signals for a given collection geometry. However, this limits the simulation to basic shapes for which an analytical model is available. A more general approach is to decompose the scene into a collection of facets. A facet is defined as a planar scattering surface of finite extent [Hunter 2006]. These are typically triangular or rectangular in nature, with triangular facets being useful because any surface can be decomposed into a collection of triangles [Hughes et al. 2013]. The response of each facet to a transmitted wave is modelled using the

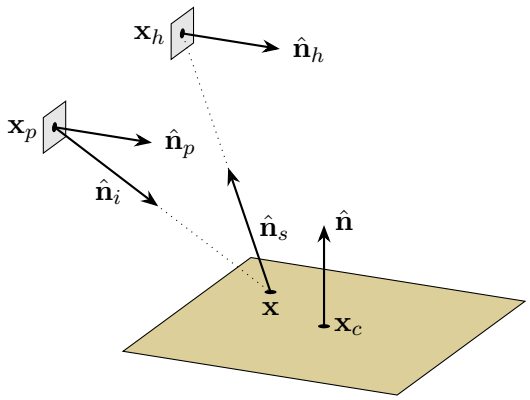


Figure 4.1: The scattering geometry for a facet centred at $\mathbf{x}_c$. A projector at $\mathbf{x}_p$ transmits acoustic energy. Some of this energy is incident upon point $\mathbf{x}$ and is scattered; a portion of this scattered energy is received by a hydrophone at $\mathbf{x}_h$. Unit normals describe the orientation of the facet, the projector, and the receiver as well as the direction of the incident and scattered energy.

properties of the underlying object, and the responses from all the facets in the scene coherently combined to form the overall dataset. Two types of facet can be used: a smooth facet which represents a uniform surface, or a rough facet which statistically models a rough surface.

A rough facet scattering model and a SAS simulator implementing this model was developed by Hunter [2006]. This chapter presents some modifications that were necessary to this simulator to allow the generation of repeat-pass SAS data. This includes the development of a model to simulate a set level of decorrelation between passes. The facet scattering model is outlined in Section 4.1. The evaluation of the beampattern of individual facets is then covered in Section 4.2. Finally, Section 4.3 presents the method developed for simulating decorrelated facets. A number of results from the simulator are given in Chapter 5.

4.1 SCATTERING FROM A FACET

The geometry of acoustic energy scattering from a facet is shown in Figure 4.1. A facet, shown here as rectangular although it may be any shape, is centred at a position $\mathbf{x}_c$ in the scene and has a normal vector $\hat{\mathbf{n}}$ describing its orientation. The projector (the transmitting transducer) centred at $\mathbf{x}_p$ insonifies the scene, and some of this energy is incident upon each point $\mathbf{x}$ in the scene. The direction between the projector and the point is denoted by the incident unit normal vector $\hat{\mathbf{n}}_i$. This energy is then scattered from the point with a portion being observed by the hydrophone (the receiving transducer) centred at $\mathbf{x}_h$. The direction of this energy is represented by the scattering unit normal vector $\hat{\mathbf{n}}_s$. The orientation of the projector and hydrophone are described by the normal vectors $\hat{\mathbf{n}}_p$ and $\hat{\mathbf{n}}_h$ respectively.

The modelling of acoustic scattering is a complex subject. This section introduces the models used in the SAS simulator, starting with the acoustic field radiated by the projector in Section 4.1.1. The field scattered by a facet is then outlined in Section 4.1.2. The extension of this to rough facets is given in Section 4.1.3. For a more complete derivation and discussion of these models, the reader is referred to the work of Hunter [2006].

4.1.1 Acoustic field radiated by the projector

The Kirchhoff-Helmholtz integral [Kirchhoff 1882] uses the values of a field ψ on a closed surface Σ and a set of boundary conditions to yield the value of the field $\psi(\mathbf{p})$ at some point $\mathbf{p}$. For a vibrating plate-style projector mounted in a supporting structure, the acoustic field ψ_i incident on the point $\mathbf{x}$ of the facet at a frequency f can be obtained by applying the Kirchhoff-Helmholtz integral as [Hunter 2006]

$$\psi_i(\mathbf{x}, f) = S(f) \iint_{\Sigma} \frac{\exp(-\mathrm{j}kr')}{4\pi r'} \left[\mathrm{j}k + \left(\mathrm{j}k + \frac{1}{r'} \right) \cos(\mathbf{r}', \hat{\mathbf{n}}'_p) \right] \mathrm{d}\mathbf{x}'_p, \tag{4.1}$$

where $S(f)$ is the signal spectrum transmitted by the projector, the surface Σ is the face of the projector, $k = 2\pi f/c$ is the transmitted wavenumber, c is the speed of wave propagation, $\mathbf{r}'$ is the vector between $\mathbf{x}$ and the point of integration $\mathbf{x}'_p$, r' is the corresponding range $|\mathbf{r}'|$, $\hat{\mathbf{n}}'_p$ is the unit normal of the projector at the point of integration, and $\cos(\mathbf{a}, \mathbf{b})$ is the *direction cosine* or cosine of the angle between the vectors $\mathbf{a}$ and $\mathbf{b}$, given by

$$\cos(\mathbf{a}, \mathbf{b}) = \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}||\mathbf{b}|}. \tag{4.2}$$

Although the expression (4.1) is valid for a point at any distance from the projector, it is inefficient to evaluate repeatedly in a simulator. If $\mathbf{x}$ is in the far field of the projector, then the Fraunhofer approximation (see Appendix B) can be used to simplify the expression. An aperture function, $a_p(\mathbf{p})$, can be defined in the plane containing the projector; this describes the shape of the projector. For the projector illustrated in Figure 4.1 the aperture function would be a two-dimensional rectangular function. Applying the Fraunhofer approximation to (4.1) gives the incident field as [Hunter 2006]

$$\psi_i(\mathbf{x}, f) = \mathrm{j}k \frac{1 + \cos(\hat{\mathbf{r}}_0, \hat{\mathbf{n}}_p)}{2} \; S(f) \frac{\exp(-\mathrm{j}kr_0)}{4\pi r_0} A_p\left(\frac{f}{c} \hat{\mathbf{r}}_0 \right), \tag{4.3}$$

where $\mathbf{r}_0$ is the vector between the centre of the projector $\mathbf{x}_p$ and $\mathbf{x}$, r_0 is the corresponding range $|\mathbf{r}_0|$, $\hat{\mathbf{r}}_0$ is the unit vector $\mathbf{r}_0/r_0$, the $(1 + \cos(\hat{\mathbf{r}}_0, \hat{\mathbf{n}}_p))/2$ term is known

as the obliquity factor, and

$$A_p(\mathbf{u}) = \iint_{-\infty}^{\infty} a_p(\mathbf{x}'_p) \exp(-j2\pi \mathbf{x}'_p \cdot \mathbf{u}) \, \mathrm{d}\mathbf{x}'_p \tag{4.4}$$

is the Fourier transform of the projector aperture function evaluated in the plane containing the projector. For simple aperture functions, the Fourier transform can be calculated analytically, while the Fast Fourier Transform algorithm [Bracewell 2000] can be used in the general case. The range-independent portions of (4.3) can then be separated out to yield the beampattern of the projector,[1]

$$B_p(\hat{\mathbf{r}}_0, f) = \mathrm{j}k(1 + \cos(\hat{\mathbf{r}}_0, \hat{\mathbf{n}}_p)) A_p\left(\frac{f}{c}\hat{\mathbf{r}}_0\right). \tag{4.5}$$

This describes the effects of the shape of the projector aperture on the radiated field at a given angle from the normal of the projector. The incident field (4.3) can be rewritten in terms of this beampattern as

$$\psi_i(\mathbf{x}, f) = S(f) B_p(\hat{\mathbf{r}}_0, f) \frac{\exp(-\mathrm{j}kr_0)}{4\pi r_0}, \tag{4.6}$$

i.e., it is a combination of the transmitted signal and the shape of the aperture with magnitude scaling and a phase shift based on the range from the projector.

4.1.2 Acoustic field scattered by the facet

A similar procedure using the Kirchhoff-Helmholtz integral combined with the Fraunhofer approximation can be used to obtain an expression for the acoustic field ψ_s scattered from the facet. The application of the Fraunhofer approximation means that the field from all points within the facet is approximated by the field from its centre point. This simplifies the scattering geometry as illustrated in Figure 4.2.

The facet can be described by an aperture function, $a(\mathbf{p})$, with the corresponding Fourier transform

$$A(\mathbf{u}) = \iint_{-\infty}^{\infty} a(\mathbf{x}') \exp(-\mathrm{j}2\pi \mathbf{x}' \cdot \mathbf{u}) \, \mathrm{d}\mathbf{x}' \tag{4.7}$$

evaluated in the plane containing the facet. The beampattern of the facet is given by

$$B(\hat{\mathbf{n}}_i, \hat{\mathbf{n}}_s, f) = \mathrm{j}\frac{2\pi R}{\lambda} \times \frac{(\hat{\mathbf{n}}_s + \hat{\mathbf{n}}_r) \cdot \hat{\mathbf{n}}}{2} \times A\left(\frac{f}{c}(\hat{\mathbf{n}}_r - \hat{\mathbf{n}}_s)\right), \tag{4.8}$$

where $((\hat{\mathbf{n}}_s + \hat{\mathbf{n}}_r) \cdot \hat{\mathbf{n}})/2$ is the facet obliquity factor, R is the reflection coefficient of the

[1] Note that although $\hat{\mathbf{r}}_0$ appears in this expression, it is a unit vector used solely for the direction to the target point rather than the range.

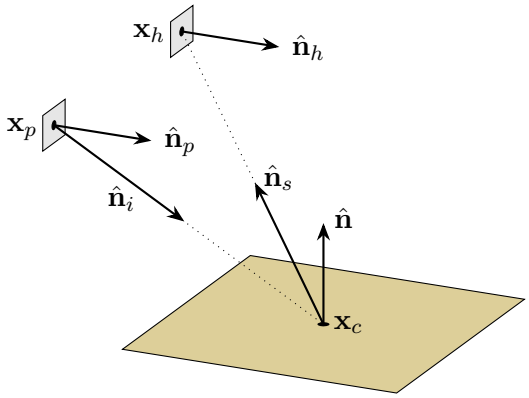


Figure 4.2: The scattering geometry for the facet after the Fraunhofer approximation has been applied. The scattered field from all points within the facet is approximated by the field from the centre of the facet.

facet and $\hat{\mathbf{n}}_r$ is the direction of reflection from the facet; both of these can be calculated from the properties of the facet[2] and the direction of the incident field $\hat{\mathbf{n}}_i$. The overall scattered field is then

$$\psi_s(\mathbf{x}_h, \mathbf{x}_p, f) = S(f) B_p(\hat{\mathbf{n}}_i, f) B(\hat{\mathbf{n}}_i, \hat{\mathbf{n}}_s, f) \frac{\exp(-\mathrm{j}2\pi f(r_i + r_s)/c)}{(4\pi)^2 r_i r_s}, \tag{4.9}$$

where r_i and r_s are the incident and scattering ranges, i.e., the ranges between the centre of the facet and the projector and hydrophone respectively.

4.1.3 Scattering from rough facets

While it is possible to model a rough surface by dividing it into a collection of smooth facets, this is computationally inefficient due to the large number of facets that would be required. A rough facet extends the smooth facet model from the previous section by considering a height function $h(\mathbf{x})$ in addition to the aperture function $a(\mathbf{x})$. The roughness statistics of the height function determine the statistics of the facet beampattern, and from this a particular realisation of the beampattern can be created. This allows a rough surface to be simulated with a greatly reduced number of facets compared to implementing it with smooth facets. Note that this assumes the roughness is unresolvable by the sonar (i.e., is smaller than its resolution cell) and thus able to be described statistically. Larger scale roughness needs to be modelled by a collection of facets.

The realisation of the beampattern for a given rough facet can be directly generated from the Fourier transform of the aperture function with a phase shift based upon the realisation of the height function for that facet [Hunter 2006]. However, this is computationally inefficient as it requires a Fourier transform at each frequency and each

[2]The facet is assumed to be on the boundary of two media, e.g., water and sand, and the reflection properties are based upon the acoustic impedance between the mediums.

set of projector and hydrophone positions. Instead, the beampattern can be separated into its specular and diffuse[3] components. The specular component is deterministic, and is able to be calculated in a similar fashion to the smooth facet beampattern. The diffuse component is the result of a random process and is difficult to directly realise. The approach used in the simulator is to take a complex Gaussian-distributed white noise process and apply a filter based upon the first- and second-order statistics of the facet roughness [Hunter 2006].

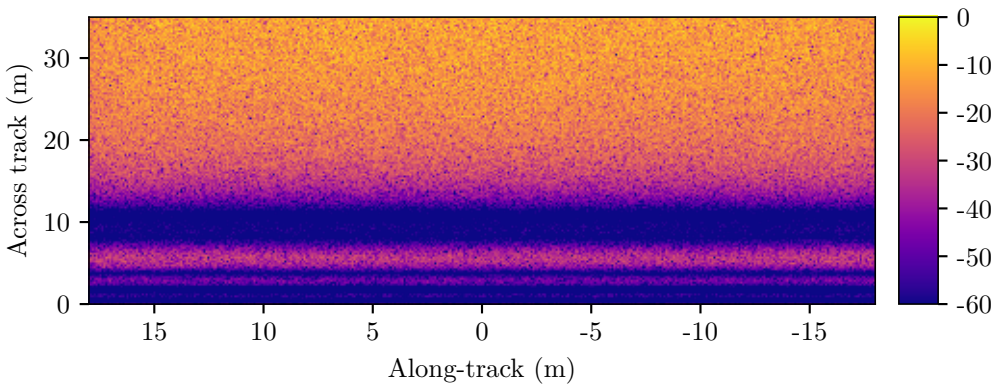
4.2 ROUGH FACET BEAMPATTERN EVALUATION

Generating noise to filter to form the diffuse component of a rough facet beampattern sounds like a simple step. However, if it is not done in a precisely repeatable fashion, then repeat-pass simulation is not possible. This was illustrated in some early tests performed with the simulator. A blank seafloor 10 m below the sonar was modelled by a set of facets, each of which was 16.7 cm by 16.7 cm in extent. Two passes on straight tracks over this scene were simulated with a 1 m difference in the across-track component of each path. The magnitude of the reconstructed imagery from the primary pass is shown in Figure 4.3a. The areas of low energy at close across track positions are a result of these areas being out of the main lobe of the beampatterns of the transducers; the stripes show where some sidelobes exist. The area within the main lobe shows a speckle image as would be expected from a coherent imaging technique.

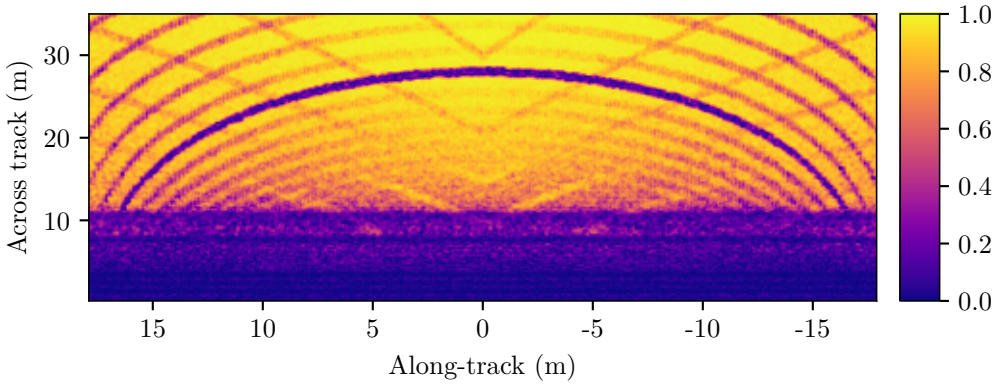
With the scene being identical and only a small amount of baseline decorrelation affecting the results, it would be expected to achieve a high coherence between the passes (apart from the areas with low beampattern energy where the image is dominated by noise). The coherence was estimated with a sliding window 20 cm by 50 cm in extent, and the resulting magnitude is shown in Figure 4.3b. Apart from the sidelobe areas, there are a number of loci where the coherence is reduced or negligible. Some of these are in the form of radial lines extending from the centre of the image, while others are arcs around the centre. The same patterns can be seen in the coherence phase image of Figure 4.3c. The interferometric phase of the two passes would be expected to be close to zero; the regions of low coherence magnitude appear as areas of increased phase difference (and in the case of the central arc, random phase).

The nature of these errors indicates a loss of coherence at certain ranges from the sonar. To test this theory, a series of simulations with single facets were carried out. In each trial, a 25 cm by 25 cm facet was placed at a different position across track. The same pair of passes as the previous simulations (10 m above the facet, and a 1 m offset between passes) were carried out and the imagery reconstructed. For each of the facet positions, 100 trials with different facet beampattern realisations were performed. For

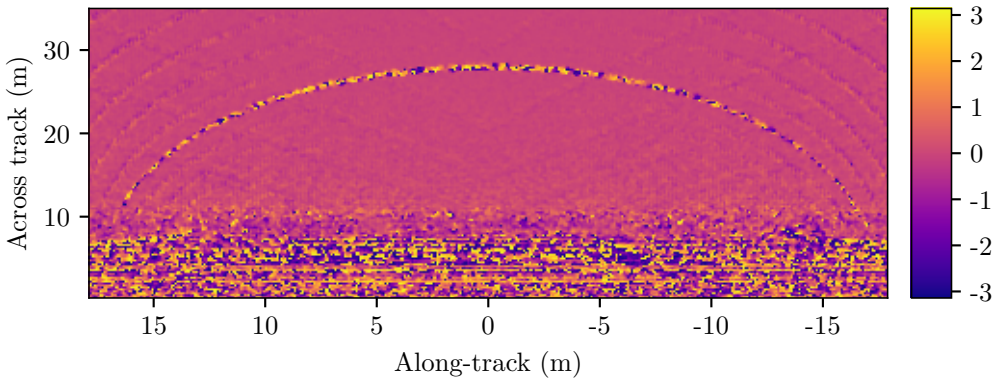
[3]Specular reflection is a reflection from a smooth surface at a definite angle. Diffuse reflection is where the energy is scattered in all directions due to a rough surface.



(a) Reconstructed magnitude image from the primary pass. The colour scale is in decibels relative to the highest magnitude in the image. The low energy at low across-track values is because those areas are out of the main lobe of the transducers.



(b) Magnitude of the coherence between passes.



(c) Phase (in radians) of the coherence between passes.

Figure 4.3: An illustration of the effects on repeat-pass simulations of incorrect noise generation for rough facets. The magnitude of the reconstructed image for the primary pass over a blank seafloor is shown in (a). A repeat pass was made with a 1 m offset in the across-track direction, and the magnitude and phase of the coherence between passes are shown in (b) and (c) respectively. The incorrect noise generation has led to loci of low coherence throughout the image.

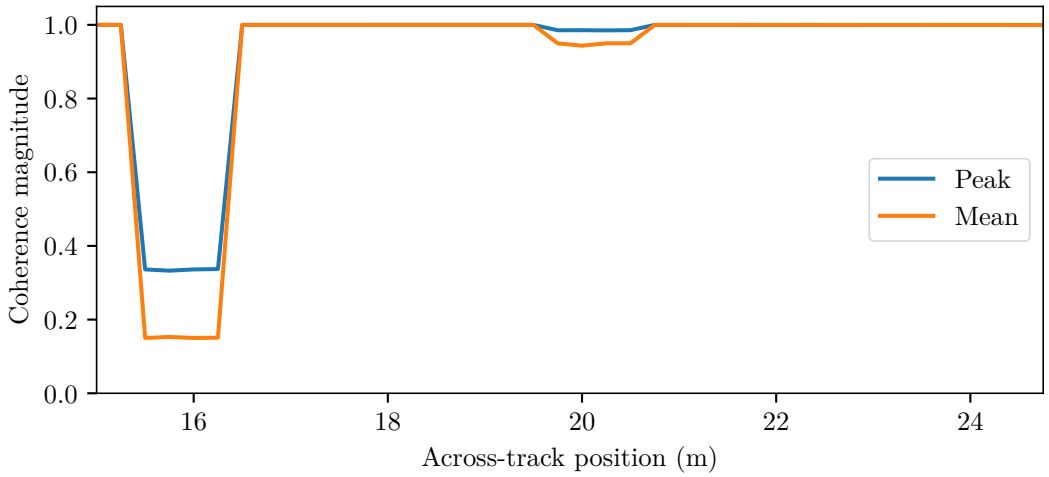


Figure 4.4: Confirmation of a problem with the beampattern noise generation. A 25 cm by 25 cm facet was placed at a number of across-track positions and 100 simulation with different beampattern realisations were performed at each position. The ensemble coherence was calculated for the reconstructed pixels within the facet boundaries, and the plot shows the peak and mean coherence magnitude observed at each across-track position. The coherence should be uniformly high for all across-track positions, but some positions show a significant loss of coherence.

each of the pixels in the reconstructed image that lay within the facet boundary, the ensemble coherence over the 100 trials was calculated. Figure 4.4 shows both the mean and peak ensemble coherence magnitude as a function of the across-track position of the facet centre. Although many of the positions result in a high coherence, there are noticeable dips at some positions. When combined across a whole scene, these could combine to form the loci previously observed.

This was eventually attributed to the way in which the noise for the diffuse component of the facet beampattern was generated. Each facet was assigned a seed for a pseudorandom number generator. The simulator then looked at the range of spatial frequencies required for the facet in the current simulation, rounded this range up to a whole number, and generated the appropriate number of random noise samples. For the repeat pass, the same seed was given to each facet; in many cases this resulted in the same set of noise samples being generated. However, due to the differing geometry (in this case, the 1 m offset between the two sonar tracks) at some positions the range of spatial frequencies was rounded to different endpoints. Although the same sequence of random numbers was generated, they were assigned to a different set of spatial frequencies and thus the resulting beampatterns were different.

The cause of the problem was difficult to find but the solution was simple. Knowing the maximum signal frequency that was to be used in all the simulations, the maximum spatial frequencies that would be needed for each facet can be calculated. This is

stored along with the seed for the pseudorandom number generator, and thus the same sequence of numbers can be generated and assigned to the same spatial frequencies in each simulation. A repeat of the single facet test showed that the ensemble coherence was now one for all across track positions. Repeating the bland seafloor simulations resulted in the repeat-pass coherence magnitude and phase shown in Figure 4.5a and Figure 4.5b respectively. Aside from the areas where there is low incident energy (and thus high noise), the coherence magnitude is one and the phase is zero throughout the scene as would be expected for the given scenario.

4.3 SIMULATING DECORRELATED FACETS

In order to model the temporal decorrelation of a scene it is necessary to be able to simulate a facet with a given level of decorrelation between passes. For two rough facets as used in the simulator, the level of correlation is controlled by the correlation of the noise sequences used to generate the realisation of their beampatterns. If the noise used on the repeat pass is not the same as used for the primary pass, but rather is a different noise realisation which has some degree of correlation with the first, then the resulting beampatterns will have the same degree of correlation.

It is straightforward to generate correlated complex Gaussian-distributed white noise as needed to generate the beampattern realisations. If $\mathbf{x}$ is a complex, zero-mean random vector then its covariance matrix is defined as

$$\operatorname{cov}(\mathbf{x}) = \mathrm{E}\big[\mathbf{x}\mathbf{x}^{\mathrm{H}}\big], \tag{4.10}$$

where $\mathbf{x}^{\mathrm{H}}$ is the Hermitian or conjugate transpose of $\mathbf{x}$. Suppose this random vector is multiplied by a matrix $\mathbf{L}$ to give a second random vector,

$$\mathbf{y} = \mathbf{L}\mathbf{x}. \tag{4.11}$$

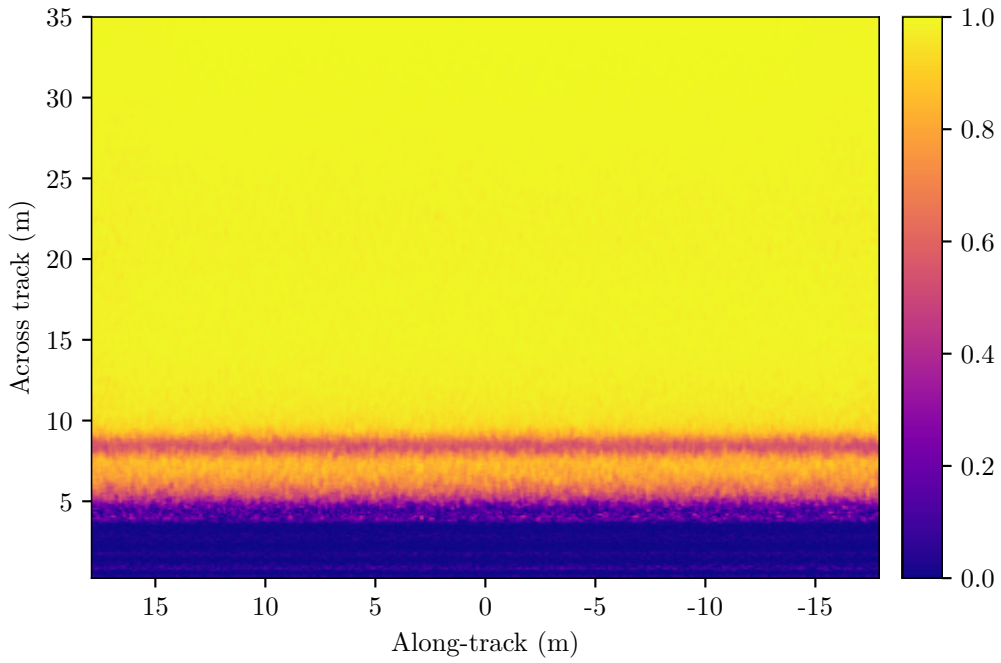
The covariance matrix of this second vector is then

$$\operatorname{cov}(\mathbf{y}) = \mathbf{L}\operatorname{cov}(\mathbf{x})\mathbf{L}^{\mathrm{H}}. \tag{4.12}$$

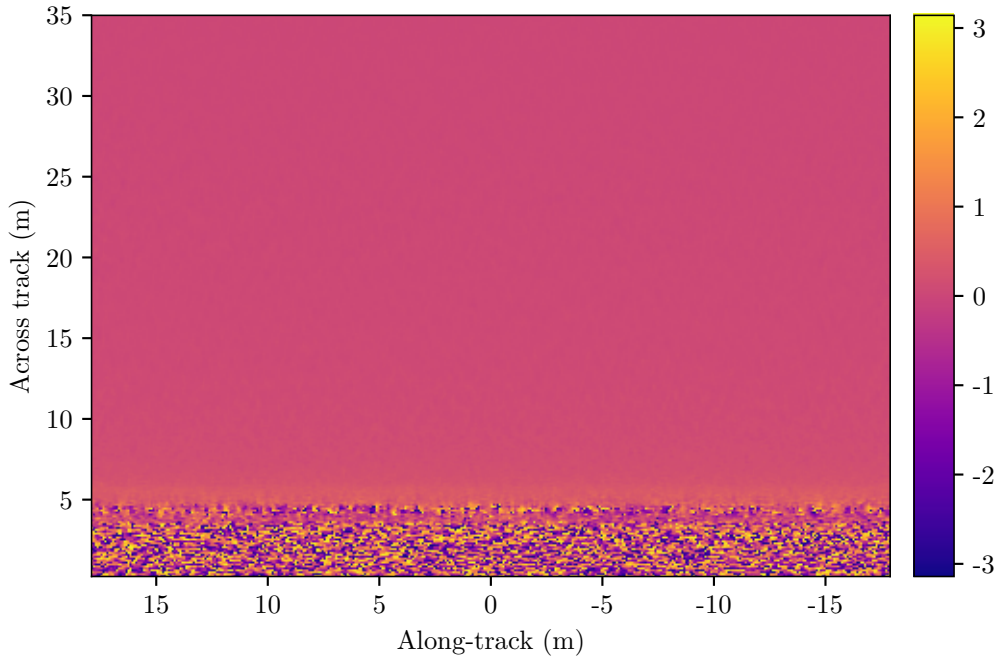
If the original vector $\mathbf{x}$ is sampled from a complex Gaussian distribution with zero mean and unit variance, then it has the identity matrix $\mathbf{I}$ for its covariance matrix. This means that the covariance matrix of $\mathbf{y}$ is

$$\operatorname{cov}(\mathbf{y}) = \mathbf{L}\mathbf{L}^{\mathrm{H}}, \tag{4.13}$$

i.e., the matrix $\mathbf{L}$ is the Cholesky decomposition [Benoît 1924; Cholesky 1910] of the covariance matrix of $\mathbf{y}$. In the two variable (two pass) case, the desired covariance



(a) Magnitude of the coherence between passes.



(b) Phase (in radians) of the coherence between passes.

Figure 4.5: The (a) magnitude and (b) phase of the estimated coherence of the blank seafloor scene after the noise generation for the rough beampattern evaluation had been corrected. Both images are uniform as expected, except where the vertical beam pattern degrades the SNR at close range.

matrix is given by

$$\mathrm{cov}(\mathbf{y}) = \sigma^2 \begin{bmatrix} 1 & \rho \\ \rho^* & 1 \end{bmatrix}, \tag{4.14}$$

where σ^2 is the variance of the individual variables and ρ the coherence between them. Applying the Cholesky decomposition to this yields

$$\mathbf{L} = \sigma \begin{bmatrix} 1 & 0 \\ \rho^* & \sqrt{1 - |\rho|^2} \end{bmatrix}. \tag{4.15}$$

In other words, from two independent samples x_1 and x_2 of a zero-mean, unit variance, complex Gaussian random variable, two samples of a correlated complex Gaussian random variable with individual variances σ^2 and coherence ρ can be created as

$$\begin{aligned} y_1 &= \sigma x_1, \\ y_2 &= \sigma \rho^* x_1 + \sigma \sqrt{1 - |\rho|^2} x_2. \end{aligned} \tag{4.16}$$

Strictly speaking, the use of the Cholesky decomposition requires the covariance matrix $\mathrm{cov}(\mathbf{y})$ to be positive-definite. As it is defined in (4.14), this corresponds to the condition $|\rho| < 1$. However, inspection of the resultant decomposition (4.15) (or equivalently the linear system of equations (4.16)) shows that the decomposition is also valid for $\rho = 1$, yielding $y_1 = y_2 = \sigma x_1$.

This technique is simple to apply to the rough facets used in the simulator. Each facet can be assigned a desired coherence value ρ for the repeat pass. As shown in Figure 4.6, two separate noise realisations are generated for each facet. The first is used for the primary pass, and a linear combination of the two as per (4.16) is used for the repeat pass. This results in two beampattern realisations with the requested coherence level ρ.

4.4 DISCUSSION

With repeatable generation of the beampattern realisations for the rough facets, the simulator described in this chapter is suitable for generating repeat-pass SAS imagery. The Fraunhofer approximation has a generally negligible effect on the coherence of the images, and any effect can be reduced at the cost of increased simulation workload if required by a particular application. The following chapter presents the results of a number of simulations carried out using the temporal decorrelation method presented here.

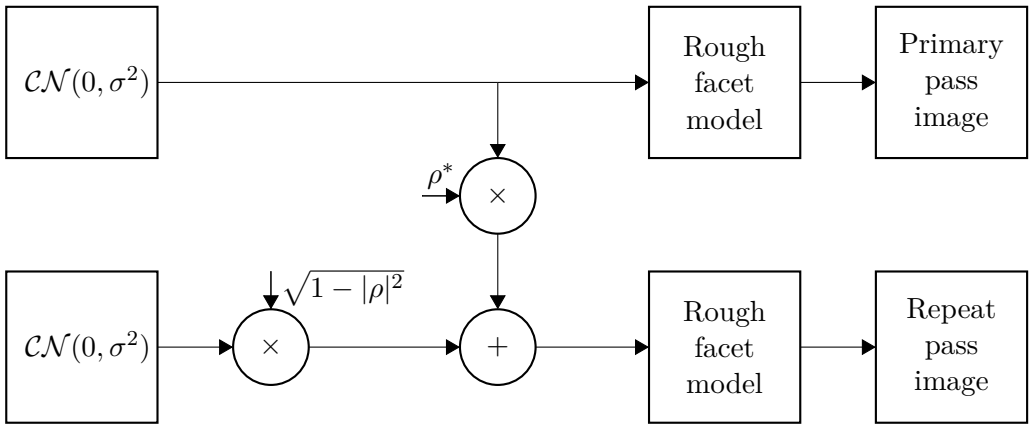


Figure 4.6: A block diagram of the proposed simulation process for a decorrelated facet. Two different noise realisations are generated. The primary pass uses the first realisation in the rough scattering model, and the repeat pass uses a linear combination of the two as per (4.16). This gives the two inputs to, and thus two outputs from, the facet model with the desired coherence ρ.

Chapter 5

RESULTS OF REPEAT-PASS SIMULATION

Chapter 4 introduced a modification to an existing rough-facet SAS simulator to model the decorrelation of the facets. This chapter evaluates its use in a variety of situations. The parameters of the sonar used in these simulations are given in Table 5.1. They were chosen to minimise other sources of error so that the majority of coherence loss in the results would be due to the modelled decorrelation of the scene. In particular, a Hamming window was applied to the transmitted signal to reduce range sidelobes, and the ping spacing was chosen to be half the smallest transmitted wavelength so that the synthetic aperture is well sampled. Both passes followed the same track so that baseline decorrelation was not a factor, no additive noise was applied, and the reconstruction process used the exact track so that footprint shift is eliminated.

In order to check the simulated imagery against the expected coherence models, statistical goodness of fit testing is required. The test used for this purpose is introduced in Section 5.1. Initial testing was done with single facets and these results are presented in Section 5.2. Simulations of a whole scene with a uniform desired coherence are then given in Section 5.3. A scene with two distinct coherence levels is shown in Section 5.4, and finally a scene with scattered patches of low coherence is presented in Section 5.5.

5.1 GOODNESS OF FIT TESTING

In goodness of fit testing, a set of observed values and a model or hypothesised distribution are compared to see to what degree they match [Massey 1951]. A number of different tests exist, some designed for testing against specific distributions while others are *distribution-free*. They are additionally classified by whether they can test against discrete or continuous distributions. In the context of this thesis, it is desired to test the fit of the magnitude and phase of the coherence measured from simulated data against the models of (3.12) and (3.16) respectively. As two complex distributions are involved, a test that is independent of the distribution is ideal; the selected Kolmogorov-Smirnov test is a continuous distribution test that meets this criteria [Facchinetti 2009].

For n independent and identically distributed observations of a random variable,

Parameter	Value
Transducer	
Width	91 mm
Height	30 mm
Tilt	26.57° down
Transmitted signal	
Type	Linear FM up-chirp
Duration	12.5 ms
Centre frequency	30 kHz
Bandwidth	20 kHz
Window	Hamming
Wavelength	37.5 mm – 75 mm
Receiver	
Ping spacing	18.75 mm
Sample rate	40 kHz
Samples per ping	2,000

Table 5.1: Parameters of the sonar used in the simulations in this chapter. The same transducer was used for both transmitting and receiving.

the empirical distribution function (EDF) $F_n(x)$ is a step function which increases by $1/n$ at each of the n data points. This means that the value of $F_n(x)$ is the fraction of observations which have a value less than x; in other words, it estimates the cumulative distribution function $F(x)$ of the underlying distribution. The Kolmogorov-Smirnov statistic[1] is a measure of the difference between the empirical and cumulative distribution functions. It is defined as [Kolmogorov 1933]

$$D_n = \sup_x |F_n(x) - F(x)|, \tag{5.1}$$

where $\sup_x$ is the supremum[2] of the set of distances. As illustrated in Figure 5.1, this can be viewed as the largest vertical difference between the two functions. The Glivenko-Cantelli theorem [Cantelli 1933; Glivenko 1933] shows that D_n converges almost surely to zero as n tends to infinity, i.e., the more observations that form the empirical distribution function, the closer it is to the cumulative distribution function.

Building on this convergence, the statistic can be used in a goodness of fit test: the Kolmogorov-Smirnov test[3]. Under the null hypothesis that the observed values come from the distribution $F(x)$, the behaviour of the Kolmogorov-Smirnov statistic is described by the Kolmogorov distribution [Kolmogorov 1933; Smirnov 1939b]. For a

[1] Sometimes referred to solely as the Kolmogorov statistic.

[2] Effectively the maximum in this situation.

[3] This section describes the one-sample Kolmogorov-Smirnov test used in this thesis. A two-sample version exists; this is used to test if two empirical distribution functions are drawn from the same distribution [Smirnov 1939a].

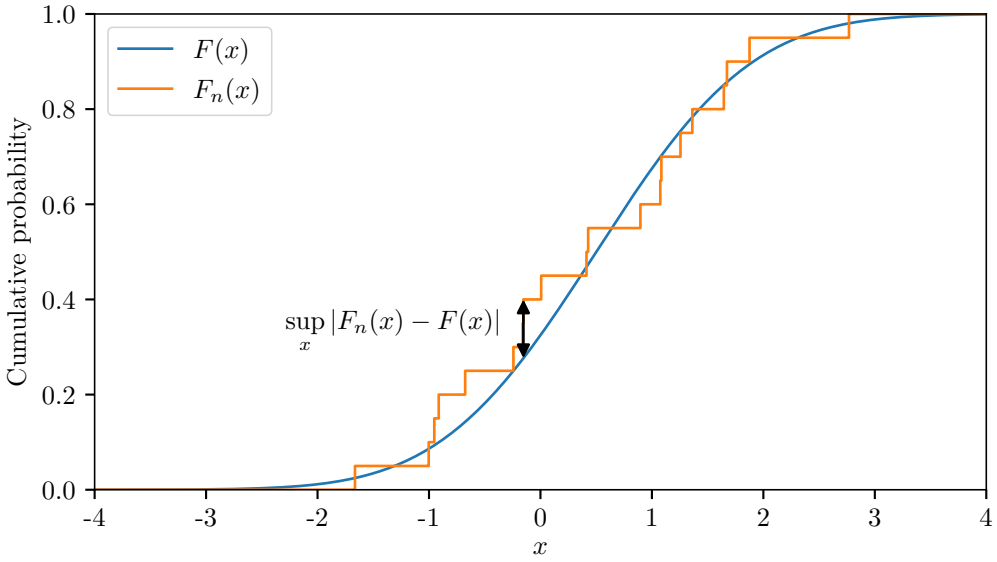


Figure 5.1: The calculation of the Kolmogorov-Smirnov statistic for a normally-distributed random variable. The blue line shows the cumulative distribution function while the orange line is the empirical distribution function from 20 observed values of the variable. The largest distance between them — marked by the arrow — is the Kolmogorov-Smirnov statistic.

significance level,[4] α, and the number of observations, n, the Kolmogorov distribution can be used to determine a critical value k. If the statistic D_n is below this critical level, then the null hypothesis that $F(x)$ describes the observed data can be accepted. Tables of the critical value for assorted values of n and α were initially published by Smirnov [1948], and for $n > 35$ the critical value can be calculated as

$$k(\alpha, n) = \sqrt{\frac{-0.5 \ln(\alpha/2)}{n}}. \tag{5.2}$$

5.2 SINGLE FACET SIMULATIONS

The initial testing of the modification to the simulator was performed with individual facets: if the facets did not perform as expected in isolation, they were unlikely to do so when grouped to form an entire scene. Each trial consisted of a facet 0.25 m × 0.25 m in size being placed 10 m below the sonar and at 25 m in the across-track domain. Two passes were simulated on identical tracks with the degree of coherence between them being specified. For each desired degree of coherence in $\delta = 0, 0.05, 0.10, \ldots, 0.90, 0.95, 0.99$, ten thousand trials were performed. Figure 5.2

[4] The probability of rejecting the null hypothesis if it is in fact true. The confidence level $\gamma = 1 - \alpha$ is the probability of accepting the null hypothesis if it is true. As a probability, α must be in the range $[0, 1]$.

shows some examples of these facets. The reconstructed imagery from the primary pass is given in Figure 5.2a, and the corresponding repeat-pass images for a variety of values of δ are in Figures 5.2b to 5.2f. The box marked in each image shows the extent of the facet. It will be noted that there is energy outside of this box; this is an artefact of the reconstruction algorithm due to limited bandwidth. The primary pass image shows a 'hook' shaped line of low energy in the bottom-left corner of the facet. As the simulated degree of coherence decreases through the repeat pass images this feature fades and disappears. In the uncorrelated $\delta = 0$ case, this has been completely replaced by different features.

For each of the trials performed, the coherence was estimated over the $0.25\,\mathrm{m} \times 0.25\,\mathrm{m}$ region of the simulated facet (this corresponds to 10×13 pixels in the reconstructed imagery). The reconstructed images are oversampled, i.e., they have a smaller resolution than the sonar system itself. This means each scatterer in the scene appears in multiple pixels in the image. When taking into account the level of oversampling and other system parameters such as the spatial windowing applied, there are $N \approx 7$ independent scatterers in the 10×13 pixel windows used for the coherence estimation.

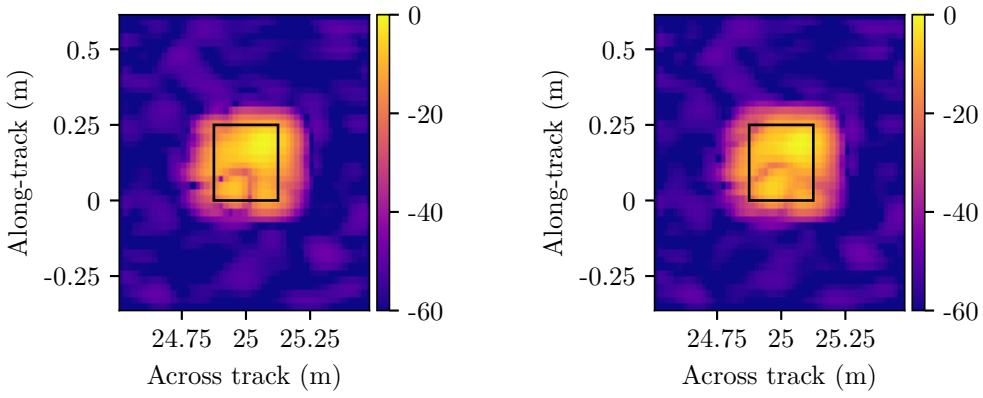
5.2.1 Coherence magnitude

Histograms of the estimated degree of coherence for $\delta = 0.95$, $\delta = 0.50$, and $\delta = 0$ are presented in Figure 5.3. Also shown on these plots are the corresponding PDFs of the model for the estimated degree of coherence given in (3.12) for $N = 7$. These show a good fit between the histogram and the model, although the lower-coherence figures exhibit more variation around the model PDF. This is especially evident in the uncorrelated $\delta = 0$ case where there are four histogram bins in a row around $d = 0.3$ which are significantly lower than the model.

The Kolmogorov-Smirnov test was applied between the 10,000 samples of the estimated degree of coherence for each value of δ and the model of (3.12). From (5.2), the critical value $k = 0.0136$ was calculated for the selected significance level $\alpha = 0.05$. The testing was performed using the SciPy library, and the results are reported in Table 5.2 along with the corresponding P values[5] returned by the library. In all but the $\delta = 0$ case the P value was higher than the significance level, and thus the null hypothesis that the model describes the data is accepted.

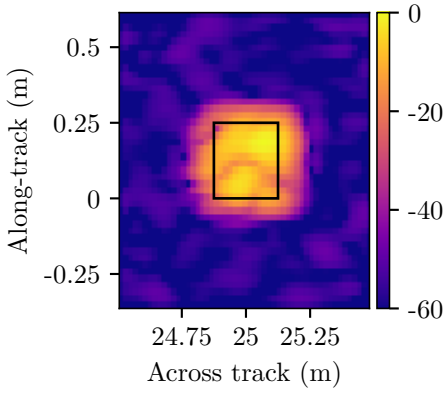
To further analyse the results for the $\delta = 0$ case, the empirical distribution function was plotted and is shown in Figure 5.4. The colour below the function indicates the Kolmogorov-Smirnov distance for that value of d, with the values below the critical value being represented by the blue colour scheme and the values above it by the red

[5]The P value is the probability that, assuming the null hypothesis is true, a more extreme data set would be observed. In this case it is the probability that a data set with a larger Kolmogorov-Smirnov statistic would be found. A low P value means the given data is at the extreme end of the scale, and the null hypothesis is rejected if the P value is lower than the significance level.

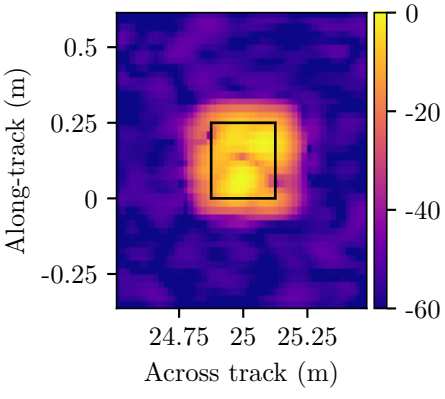


(a) Primary pass.

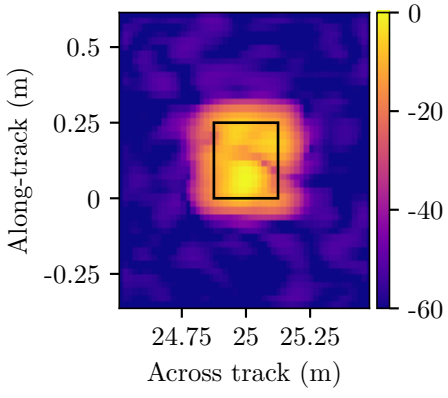
(b) Repeat pass, $\delta = 0.95$.



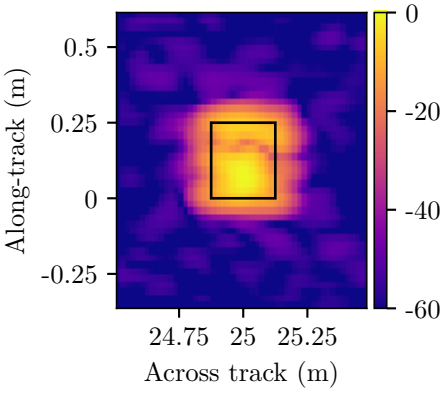
(c) Repeat pass, $\delta = 0.90$.



(d) Repeat pass, $\delta = 0.75$.

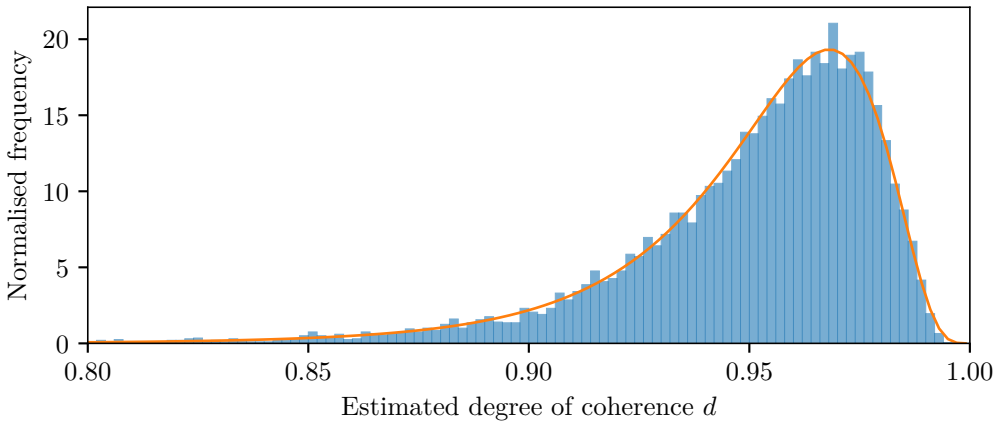


(e) Repeat pass, $\delta = 0.50$.

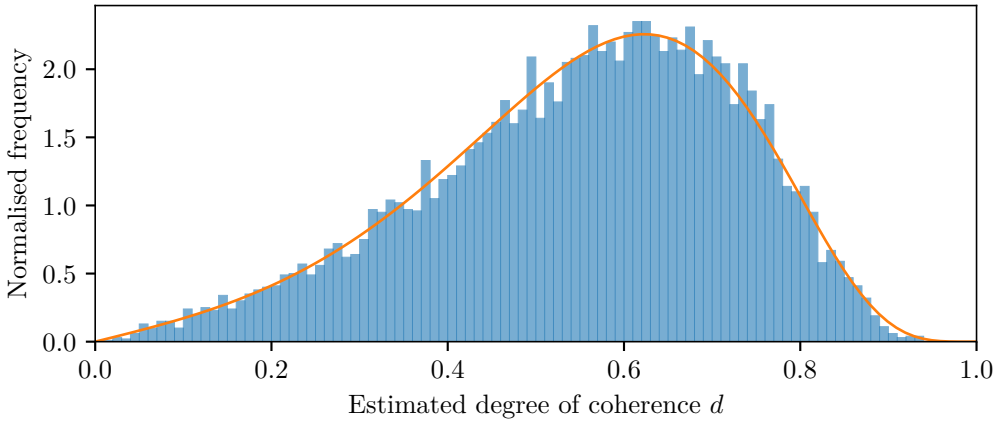


(f) Repeat pass, $\delta = 0$.

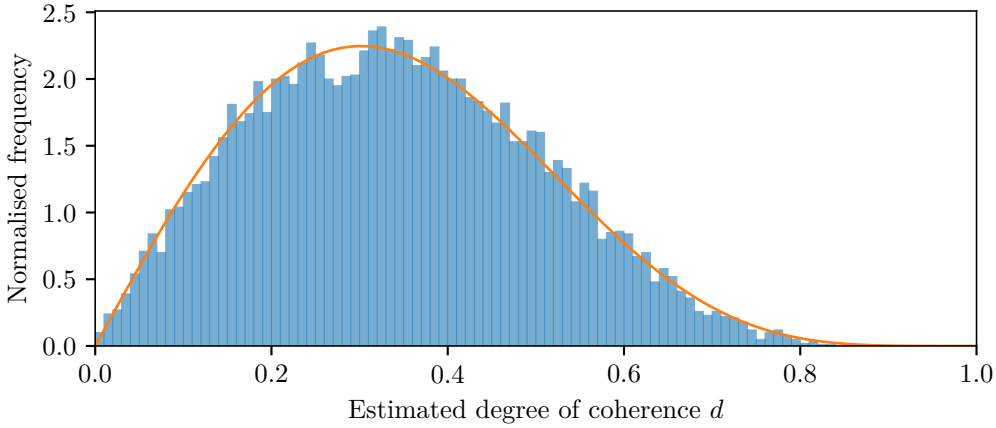
Figure 5.2: An example of simulated repeat-pass facets with varying degrees of coherence. The reconstructed primary pass is shown in (a), and (b) – (f) show the reconstructed repeat passes as the specified degree of coherence decreases. The colour scale is the magnitude of the reconstructed image in decibels relative to the highest value in the images. The marked box shows the extent of the simulated facet; any energy outside this box is an artefact of the reconstruction algorithm.



(a) $\delta = 0.95$.



(b) $\delta = 0.5$.



(c) $\delta = 0$.

Figure 5.3: Histograms of the estimated degree of coherence for single-facet simulations with various values of the true degree of coherence δ. Each histogram contains the results of 10,000 trials, and the overplotted line is the PDF of the model (3.12) assuming $N = 7$ independent scatterers in the estimation window.

δ	Kolmogorov-Smirnov statistic D_n	P value
0.00	0.014	0.03
0.05	0.013	0.08
0.10	0.010	0.22
0.15	0.010	0.27
0.20	0.010	0.22
0.25	0.011	0.16
0.30	0.010	0.22
0.35	0.010	0.32
0.40	0.010	0.24
0.45	0.009	0.43
0.50	0.009	0.41
0.55	0.010	0.29
0.60	0.010	0.27
0.65	0.011	0.21
0.70	0.012	0.12
0.75	0.011	0.19
0.80	0.012	0.13
0.85	0.010	0.23
0.90	0.011	0.18
0.95	0.010	0.27
0.99	0.009	0.35

Table 5.2: The Kolmogorov-Smirnov test results for the fit of the estimated degree of coherence measured from 10,000 single-facet simulations to the model given in (3.12). The critical value for the significance level $\alpha = 0.05$ is 0.0136.

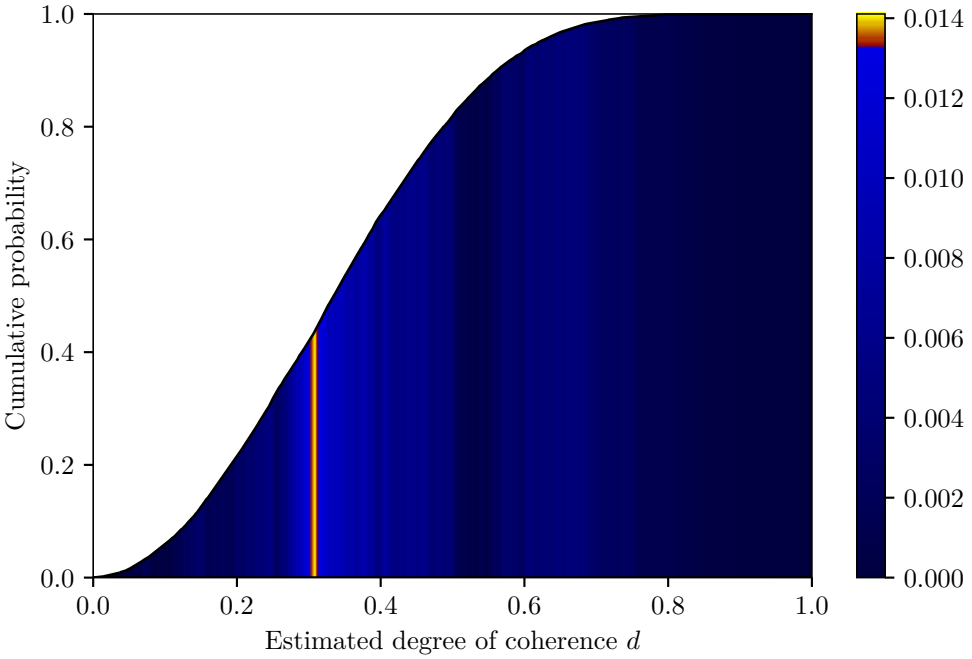


Figure 5.4: The empirical distribution function of the estimated degree of coherence d for 10,000 simulated facets with a specified degree of coherence $\delta = 0$. The colours under the plotted function indicate the Kolmogorov-Smirnov distances when comparing the empirical distribution function to the cumulative distribution function of the model (3.12). The blue colours are below the critical value of 0.0136 (for a significance level $\alpha = 0.05$), while the red and yellow colours are above it.

and yellow colour scheme. The maximum of these values is the Kolmogorov-Smirnov statistic for this data set. From the figure, it is clear that the majority of the empirical distribution function has a Kolmogorov-Smirnov distance below the critical value. The band where the distance is above the critical value corresponds to the previously-noted area in the corresponding histogram (Figure 5.3c) where several bins in a row are significantly below the PDF of the model. This suggests that this dip in the histogram is causing the rejection of the model for $\delta = 0$.

Examining the linear combination used to generate the facet realisation in (4.16), for $\delta = 0$ only the second noise realisation is used. This means the realisations for the two passes are by definition independent. Therefore it is unlikely that the simulator is causing this dip in the histogram. It is hypothesised that this is simply caused by an 'unlucky' set of realisations from the pseudo-random number generator used to create the noise input into the simulator. The trials were divided into ten subsets of 1,000 trials each, and the Kolmogorov-Smirnov test applied to each of these subsets. The results are presented in Table 5.3. Eight of these ten subsets have a P value well above the significance level, and thus the model is accepted to describe the data in those subsets. The subset of trials 3001–4000 has a marginal P value of 0.05, and the subset of trials

Subset	Kolmogorov-Smirnov statistic D_n	P value
1 – 1000	0.032	0.25
1001 – 2000	0.016	0.96
2001 – 3000	0.025	0.57
3001 – 4000	0.042	0.05
4001 – 5000	0.025	0.55
5001 – 6000	0.017	0.94
6001 – 7000	0.047	0.03
7001 – 8000	0.024	0.62
8001 – 9000	0.018	0.91
9001 – 10000	0.021	0.76

Table 5.3: The Kolmogorov-Smirnov test results for subsets of the estimated degree of coherence from single-facet simulations with $\delta = 0$. The critical value for the significance level $\alpha = 0.05$ is 0.0429.

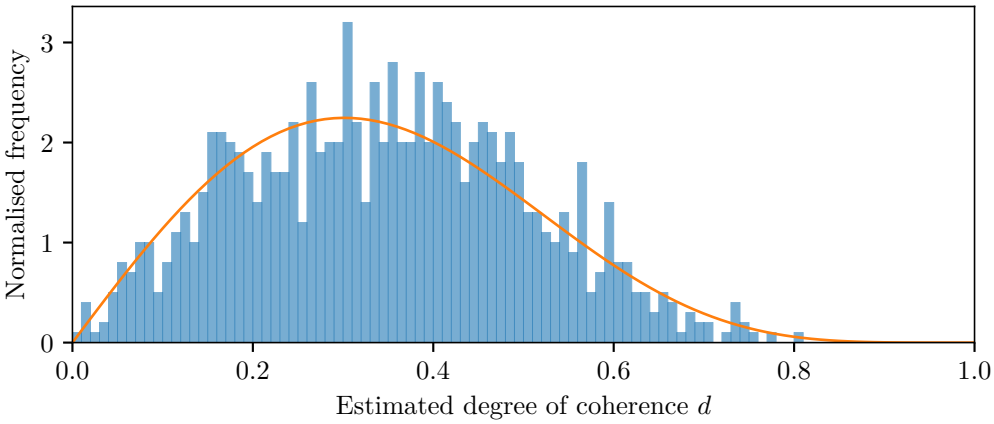
6001–7000 has a P value of 0.03, the same as the complete set of trials. The histograms of values for these subsets are shown in Figure 5.5. Both of these subset histograms show large differences from the model at values of d which are not observed in the full histogram of Figure 5.3c. This backs up the hypothesis that the difference from the model is due to the underlying random number generator (and possibly the value with which it was seeded) rather than an error in the simulation process or implementation. It is also worth noting that the model parameters may not be exact. For example, the number of independent scatterers N may be 7.05 rather than the 7 used here, causing a slight error which was exaggerated by this particular case.

5.2.2 Coherence phase

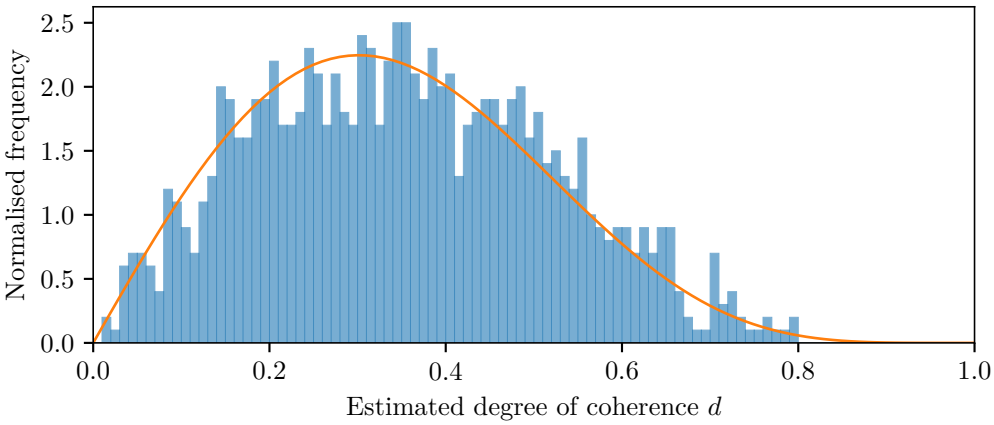
Histograms of the estimated coherence phase for $\delta = 0.95$, $\delta = 0.5$, and $\delta = 0$ are shown in Figure 5.6 along with the correponding PDFs from the phase model (3.16). These show a good fit between the values estimated from the trials and the model. To quantify this, the Kolmogorov-Smirnov test was applied with the results presented in Table 5.4. For all of the simulated δ values, the P value is above the significance level $\alpha = 0.05$, i.e., the null hypothesis that the simulated values fit the phase model is accepted.

5.3 WHOLE SCENE SIMULATIONS

Having tested the ability of the simulator to generate single decorrelated facets, the next step was to evaluate a scene consisting of multiple facets. A bland seafloor scene 10 m in across-track extent and 15 m in along-track extent was simulated. This consisted of 3,750 facets, each of which was 0.2 m × 0.2 m in size. The scene was centred 10 m below and 26 m to the side of the sonar. Identical tracks were used for both passes, and the

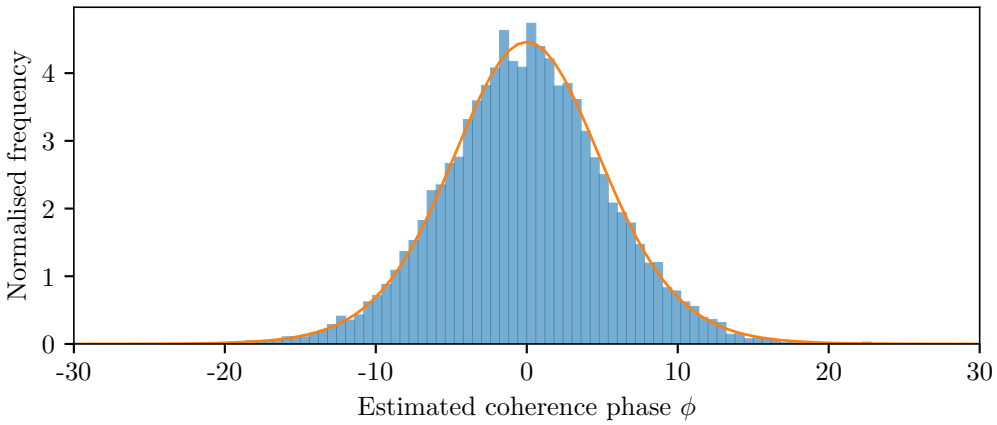


(a) Trials 3001 – 4000.

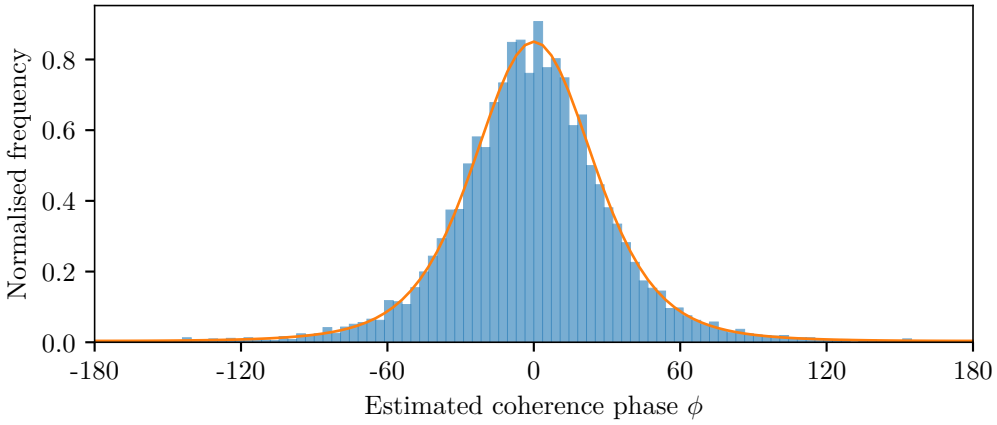


(b) Trials 6001 – 7000.

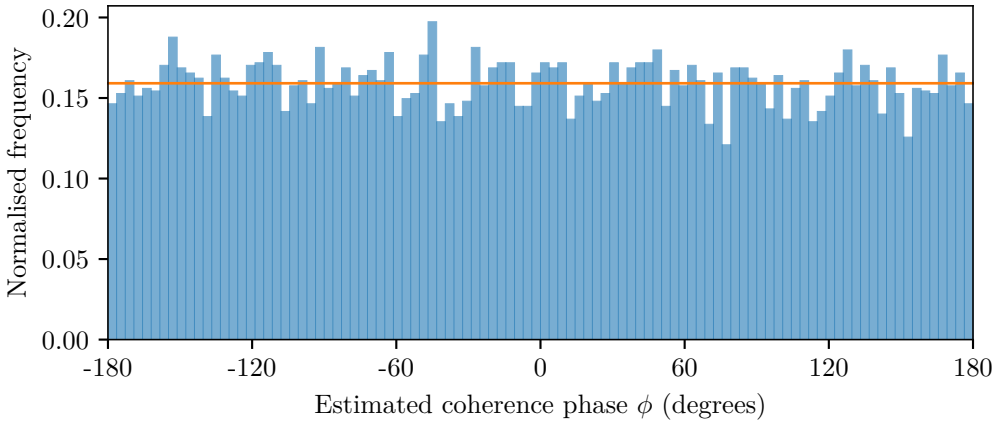
Figure 5.5: Histograms of subsets of the estimated degree of coherence for single-facet simulations when $\delta = 0$. The overplotted line is the PDF of the model (3.12) assuming $N = 7$ independent scatterers in the estimation window.



(a) $\delta = 0.95$.



(b) $\delta = 0.5$.



(c) $\delta = 0$.

Figure 5.6: Histograms of the coherence phase for single-facet simulations with various values of the true degree of coherence δ. Each histogram contains the results of 10,000 trials, and the overplotted line is the PDF of the model (3.16) assuming $N = 7$ independent scatterers in the estimation window.

δ	Kolmogorov-Smirnov statistic D_n	P value
0.00	0.009	0.38
0.05	0.011	0.16
0.10	0.011	0.16
0.15	0.011	0.15
0.20	0.010	0.26
0.25	0.010	0.22
0.30	0.011	0.18
0.35	0.011	0.19
0.40	0.012	0.14
0.45	0.012	0.14
0.50	0.012	0.12
0.55	0.012	0.10
0.60	0.012	0.11
0.65	0.012	0.09
0.70	0.013	0.08
0.75	0.012	0.10
0.80	0.012	0.11
0.85	0.012	0.11
0.90	0.012	0.12
0.95	0.012	0.12
0.99	0.012	0.10

Table 5.4: The Kolmogorov-Smirnov test results for the fit of the estimated coherence phase measured from 10,000 single-facet simulations to the model given in (3.16). The critical value for a significance level $\alpha = 0.05$ is 0.0136.

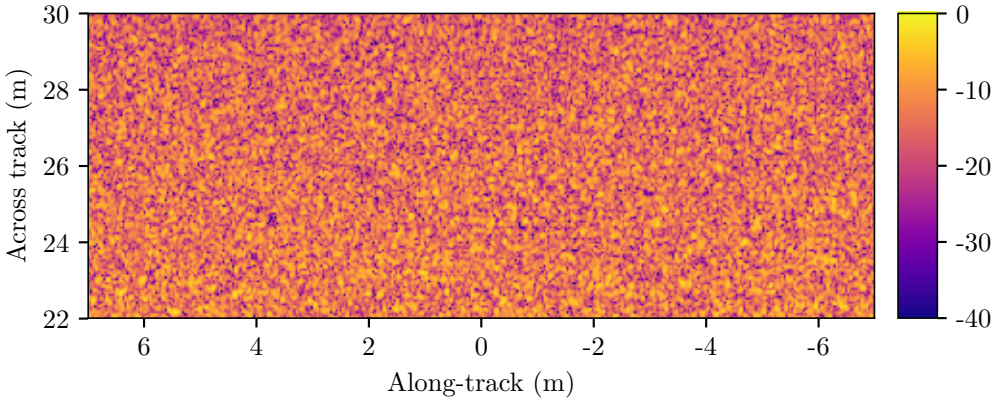
repeat-pass coherence was set to $\delta = 0.9$. The reconstructed images of the two passes are shown in Figure 5.7a and Figure 5.7b. The coherence was estimated using (3.10) over 0.19 m by 0.24 m windows and its magnitude is given in Figure 5.7c. As with the previous experiments, this corresponds to $N \approx 7$ independent scatterers per window.

Due to the sliding nature of the process, adjacent windows share the majority of the values within them. To avoid any statistical dependence, 2,475 samples of the coherence were selected from non-overlapping windows. Histograms of the magnitude and phase of these samples are plotted in Figure 5.8a and Figure 5.8b respectively. Also shown are the PDFs of the corresponding models ((3.12) for the magnitude and (3.16) for the phase) for $N = 7$ independent scatterers in the estimation window. The Kolmogorov-Smirnov test was applied to these samples with a critical value $k = 0.0273$ for the significance level $\alpha = 0.05$. The magnitude data had a statistic of 0.0230 (corresponding to a P value of 0.147) and the phase data had a statistic of 0.009 (P value = 0.988) indicating that the data was a good fit to the reference models. This confirms that multiple adjacent rough facets can be used to simulate a larger scene with a specified coherence loss.

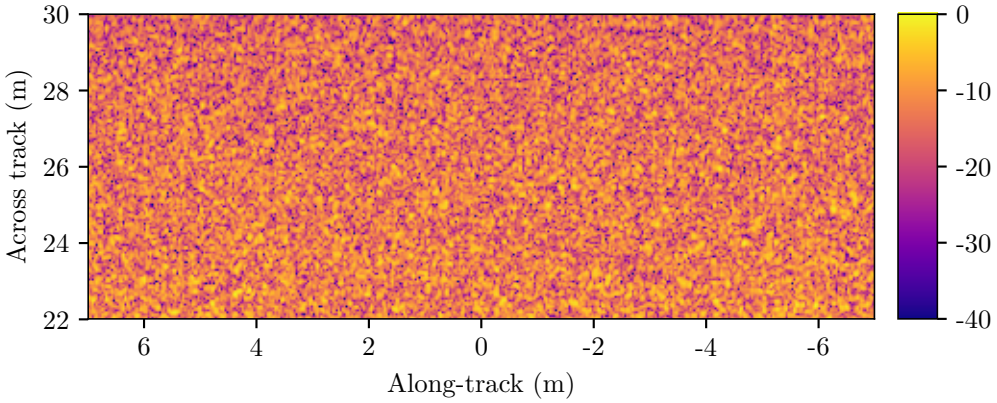
5.4 SPLIT SCENE SIMULATIONS

The next test was a simulation with two different repeat-pass coherences. The same bland seafloor scene was used, but for the repeat pass, half of the scene was set to have a coherence $\delta = 0.9$ and the other half $\delta = 1$ (i.e., no decorrelation). The reconstructed image for the primary pass is given in Figure 5.9a and the repeat pass in Figure 5.9b. The coherence was estimated over a sliding 0.19 m by 0.24 m window, and the estimated degree of coherence is shown in Figure 5.9c. This clearly shows the divide between the two areas with different coherence levels.

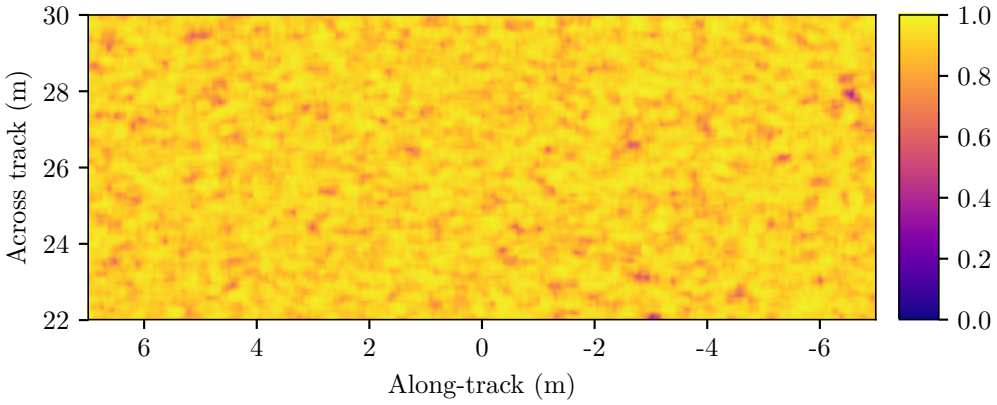
As with the previous simulation, the estimated coherences were then sampled so that all samples came from non-overlapping windows. Histograms of the estimated degree of coherence and phase are presented in Figure 5.10a and Figure 5.10b respectively. The PDFs of the corresponding models for the $\delta = 0.9$ case are also plotted. The samples from the perfectly correlated ($d = 1$ and $\phi = 0$) part of the scene are clearly distinguishable from the decorrelated area. The 1,221 samples from the decorrelated portion of the scene were used in Kolmogorov-Smirnov tests with both the magnitude and phase models. With a critical value of 0.0389 for the significance level $\alpha = 0.05$, both components of the coherence showed a good fit: the magnitude had a statistic of 0.0289 (a P value of 0.524) and the phase had a statistic of 0.0251 (a P value of 0.422). Therefore the null hypothesis can therefore be accepted, and this shows that having adjacents areas with different repeat-pass coherence values does not affect the validity of the simulation.



(a) Primary pass.

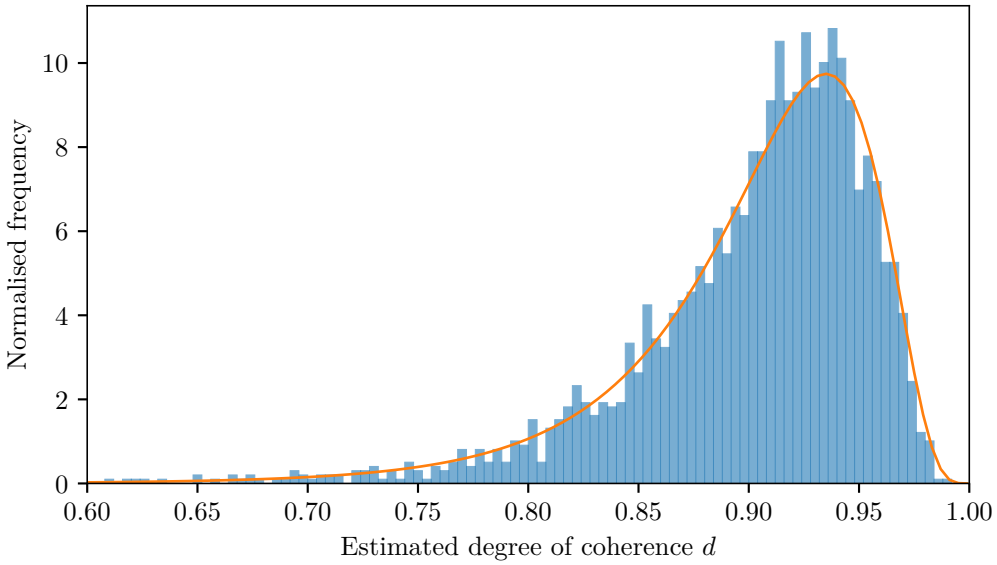


(b) Repeat pass; $\delta = 0.9$.

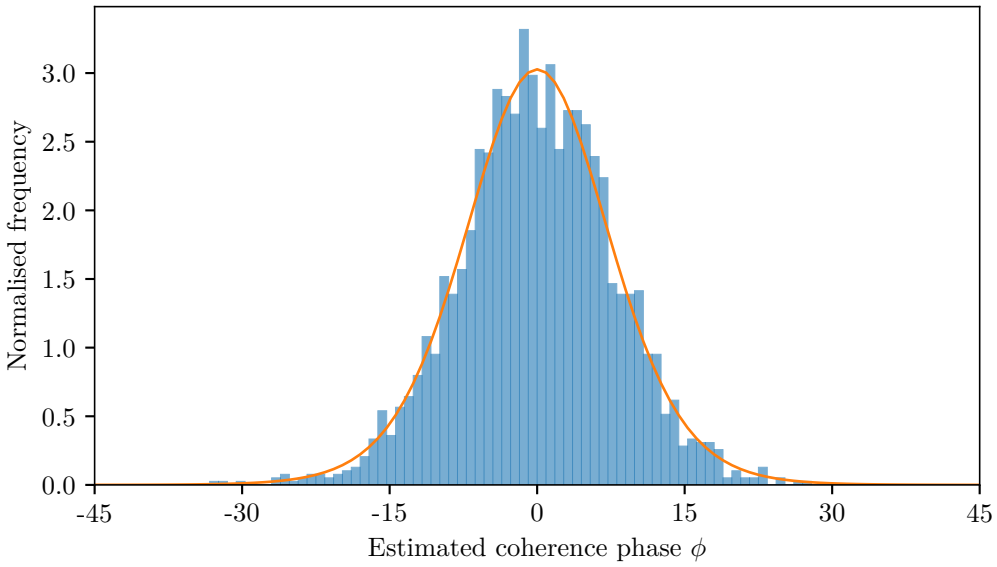


(c) Coherence map.

Figure 5.7: A repeat-pass simulation of a bland seafloor. The area shown consists of approximately 1,400 facets each 0.2 m × 0.2 m in extent. Magnitude images of the primary and repeat passes are shown in (a) and (b) respectively with the colour scales being decibels relative to the highest magnitude in the scene. The desired repeat-pass coherence was set to $\delta = 0.9$. A map of the estimated degree of coherence is shown in (c); this was estimated over a sliding 0.19 m × 0.24 m window.

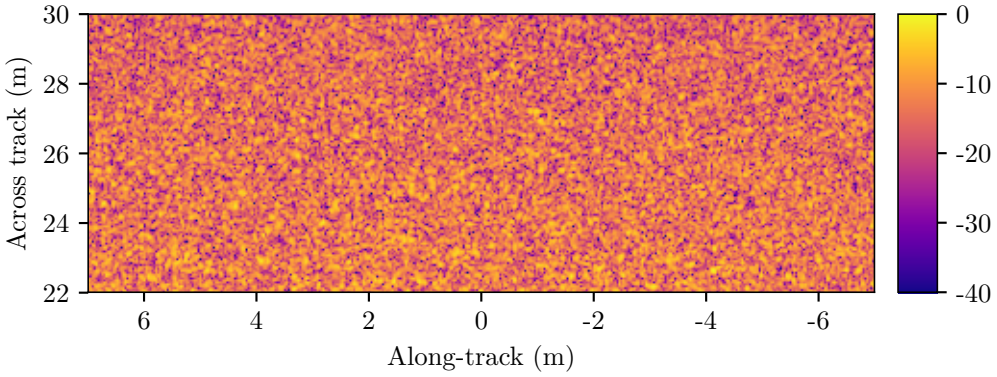


(a) Degree of coherence.

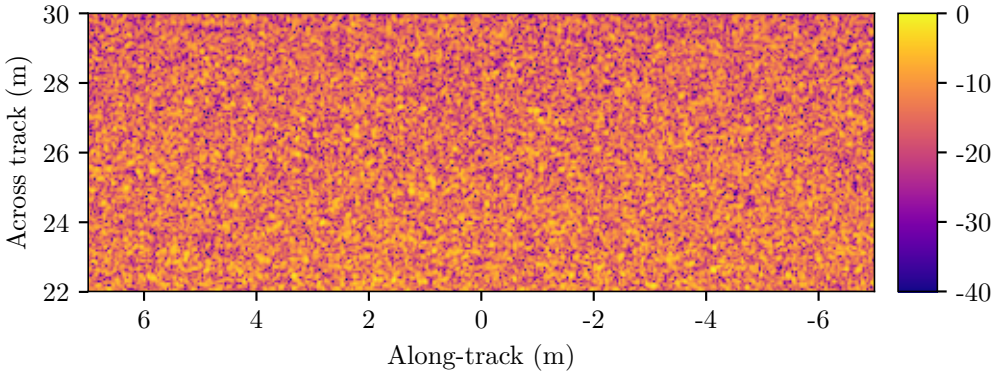


(b) Coherence phase.

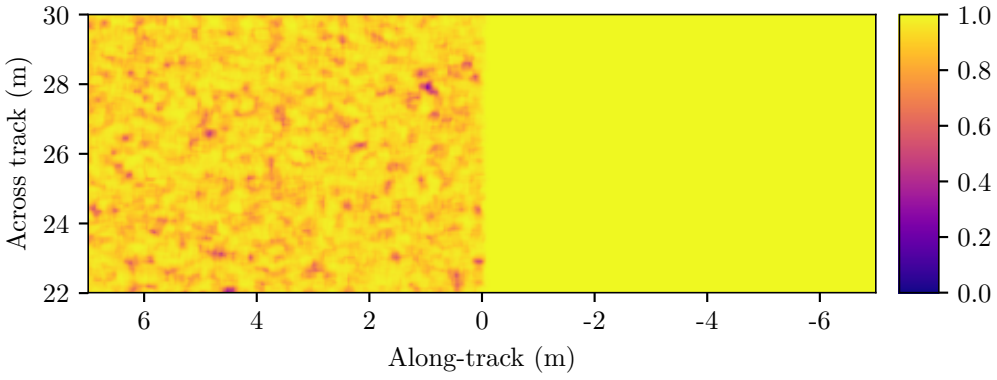
Figure 5.8: Histograms of (a) the estimated degree of coherence and (b) the estimated coherence phase from the simulated bland seafloor presented in Figure 5.7. Each histogram contains 2,475 samples selected so that they were calculated using non-overlapping windows. The overplotted lines show the PDFs from the corresponding models (3.12) and (3.16) for $N = 7$ independent scatterers in the estimation window.



(a) Primary pass.

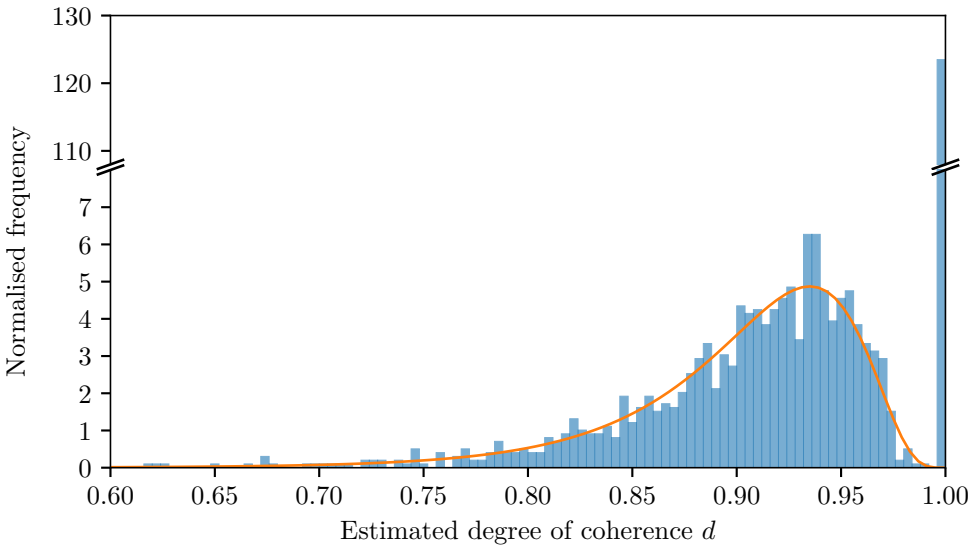


(b) Repeat pass; $\delta = 0.9$.

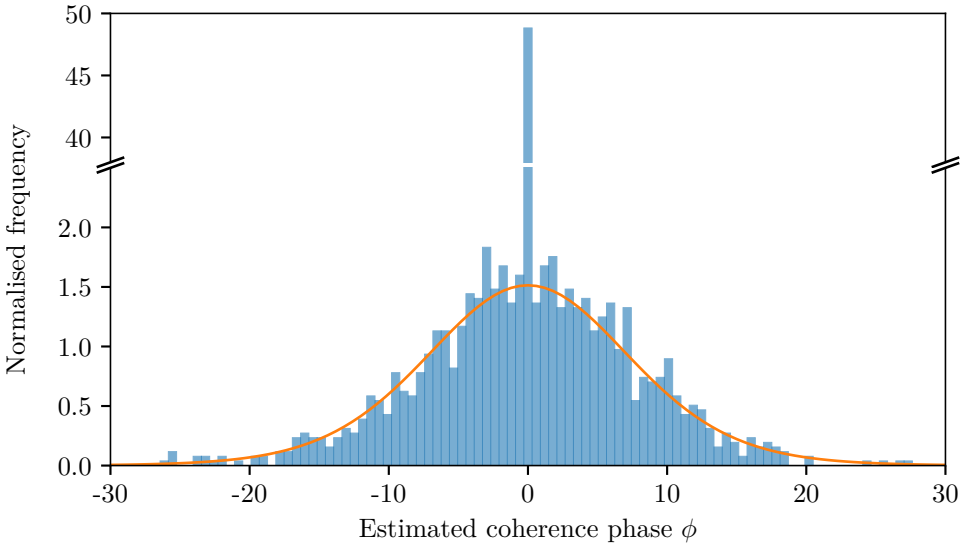


(c) Coherence map.

Figure 5.9: A repeat-pass simulation of a bland seafloor with half the scene undergoing a loss of coherence between passes. The area shown consists of approximately 1,400 facets each 0.2 m × 0.2 m in extent. Magnitude images of the primary and repeat passes are shown in (a) and (b) respectively with the colour scales being decibels relative to the highest magnitude in the scene. The desired repeat-pass coherence was set to $\delta = 0.9$ for the half of the scene in the positive along-track axis, and $\delta = 1$ (i.e., no decorrelation) for the rest of the scene. A map of the estimated degree of coherence is shown in (c); this was estimated over a sliding 0.19 m × 0.24 m window.



(a) Degree of coherence.



(b) Coherence phase.

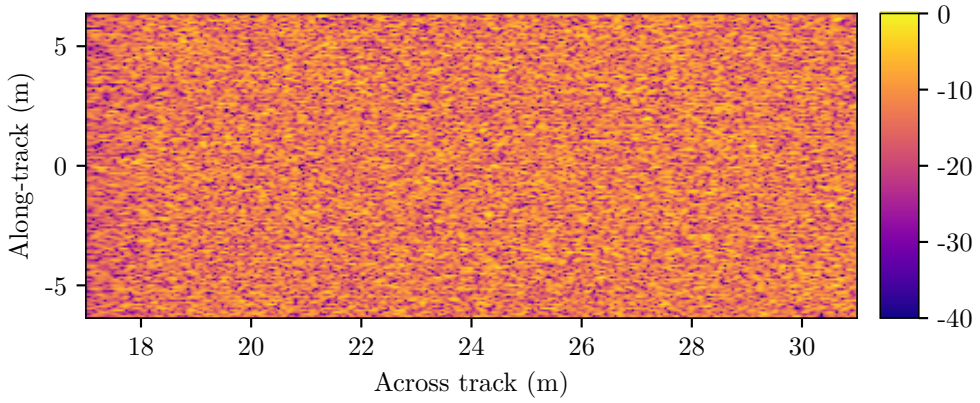
Figure 5.10: Histograms of (a) the estimated degree of coherence and (b) the estimated coherence phase from the simulation presented in Figure 5.9. As half of the scene was decorrelated between passes there are effectively two distributions within the entire scene; note the split in the y axes to accomodate this. The overlaid lines show the corresponding model PDFs for the decorrelated portion of the scene.

5.5 SCATTERED SCENE SIMULATIONS

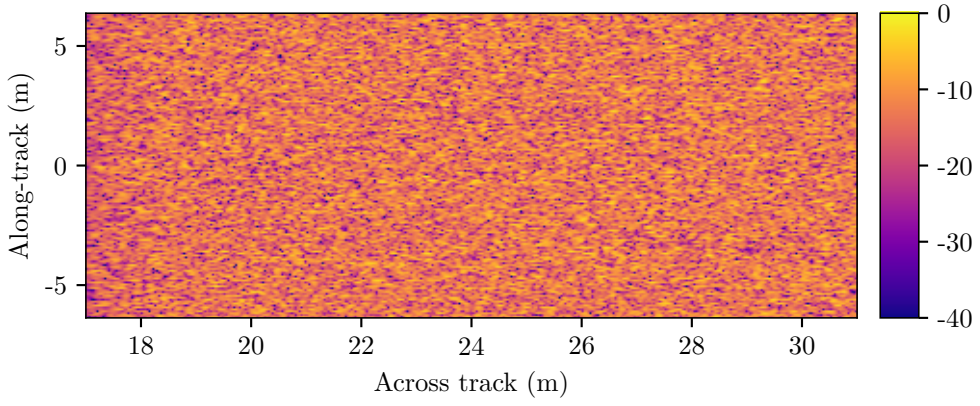
The final experiment performed to validate the changes to the simulator was to simulate a bland seafloor with changes in coherence scattered throughout the scene. The correlation map shown in Figure 2.3d (reproduced from G-Michael, Marchand, Tucker, Marston, et al. [2016]) was used as the source of the changes. The scene consisted of 71,000 5 cm by 5 cm facets, each of which was assigned a repeat-pass coherence value from the source image. Both the primary and repeat passes followed the same track 10 m above the seafloor, and the reconstructed images are shown in Figure 5.11a and Figure 5.11b respectively. The coherence was estimated over a 15 cm by 15 cm window with the estimated degree of coherence given in Figure 5.11c. It is not identical to the source image; without knowing the exact reconstruction procedure and subsequent processing applied to obtain that image it would be difficult to achieve this. However, it does show the same basic features (in particular, the track running between the corners the original authors attributed to fauna) and thus confirms the modification to the simulator is suitable for generating such scenes.

5.6 DISCUSSION

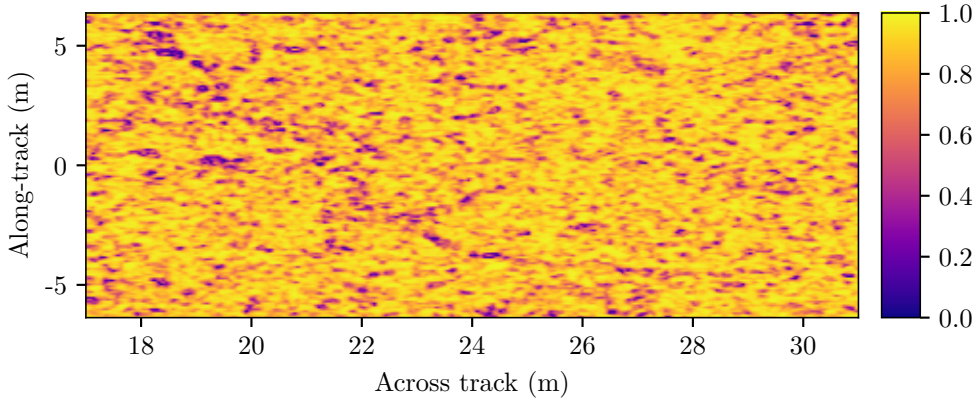
The modification to the rough facet simulator presented in Section 4.3 to simulate a decorrelated facet for repeat-pass images has been shown to generate valid data. The repeat-pass coherence of individual facets has been shown to fit standard statistical models for both the magnitude and phase. When multiple facets are combined to form a scene, the coherence over the entire scene also matches these models. The ability of this modification to generate scattered patches of low coherence based on real-world results has also been demonstrated.



(a) Primary pass.



(b) Repeat pass.



(c) Coherence map.

Figure 5.11: A repeat-pass simulation of a seafloor with scattered changes. The area shown consists of approximately 71,000 facets each 5 cm by 5 cm in extent. Magnitude images of the two passes are shown in (a) and (b) respectively; the colour scale is in decibels relative to the highest magnitude in each image. For the repeat pass the decorrelation of each facet was set using Figure 2.3d as the source. The coherence map, in this case estimated over a 15 cm × 15 cm window, is shown in (c).

Chapter 6

REGISTRATION OF REPEAT-PASS IMAGES

As detailed in Section 3.6.2, the key to change detection is the accurate registration of the imagery. This requires knowledge of the position of the sonar for all pings in order to minimise the footprint shift. The vessel towing or carrying the sonar will have some idea of its location for navigation purposes. This is recorded concurrently with the raw sonar data, providing an initial estimate of the location of the reconstructed data and therefore the common areas of the different imagery. In most cases, this will not result in a sufficiently accurate registration. At this point, it becomes necessary to perform *data-driven alignment*, i.e., to use the images themselves to improve the registration. Having estimated a correction to the navigation, there are two methods of applying it. Firstly, the repeat image itself can be warped to match the primary image. A second method, and the one used in this research, is to update the navigation data and regenerate the repeat-pass imagery with the correct track. This has the benefit of minimising the interpolation applied to the final data. As shown in Section 3.6.4.2 interpolation is a source of noise and thus a factor in the loss of coherence.

Section 6.1 looks at the issue of reconstructing the data from multiple runs onto a common grid. Section 6.2 then looks at the accuracy available with various navigation techniques and sensors. The use of navigation data for registration is evaluated for a real-world system instrumented with high-quality navigation data in Section 6.3.

A model explaining how errors in the track affect the alignment of the imagery is required in order to estimate these errors from the data. The model used in this research is introduced in Section 6.4, and the method used to estimate the registration errors between the two images in given in Section 6.5. From these registration errors, the errors in the assumed track can be estimated; two estimators are presented in Section 6.6. The parameters of a sonar system used for simulating SAS data for testing these estimators are given in Section 6.7. This data is then used to evaluate the performance of the presented registration algorithms in the presence of translation and rotation errors, and the results of this are detailed in Section 6.8 and Section 6.9 respectively.

6.1 ALIGNMENT TO A COMMON GRID

In a standard single-pass sidescan system, sonar images are reconstructed onto a co-ordinate system with the axes being the along-track axis (parallel with the direction of travel) and across-track axis (perpendicular to the direction of travel).[1] For repeat-pass applications, each image needs to use a common grid in order to allow them to be compared. Note that the choice of coordinate system is arbitrary: as long as the data from each run is reconstructed using the same coordinate system, they can be compared. For convenience, it makes sense to use a coordinate system which preserves some idea of along- and across-track axes. Section 6.1.1 presents a method of linearly approximating an arbitrary track followed by a sonar, and Section 6.1.2 discusses the selection of a coordinate system based upon these approximations and shows how the original tracks can be transformed into the new coordinate system.

6.1.1 Linear tracks

Let the position of the sonar at the p^{th} ping of the m^{th} run be described by the vector

$$\mathbf{x}_m(p) = \begin{bmatrix} x_m(p) \\ y_m(p) \\ z_m(p) \end{bmatrix}. \tag{6.1}$$

A linear approximation of the track described by these positions is defined by the centre of the track $\mathbf{c}_m$ and the displacement vector $\mathbf{d}_m$ (the displacement between subsequent pings). Estimates of these two values, $\hat{\mathbf{c}}_m$ and $\hat{\mathbf{d}}_m$ respectively, can be formed using a linear least-squares approach:

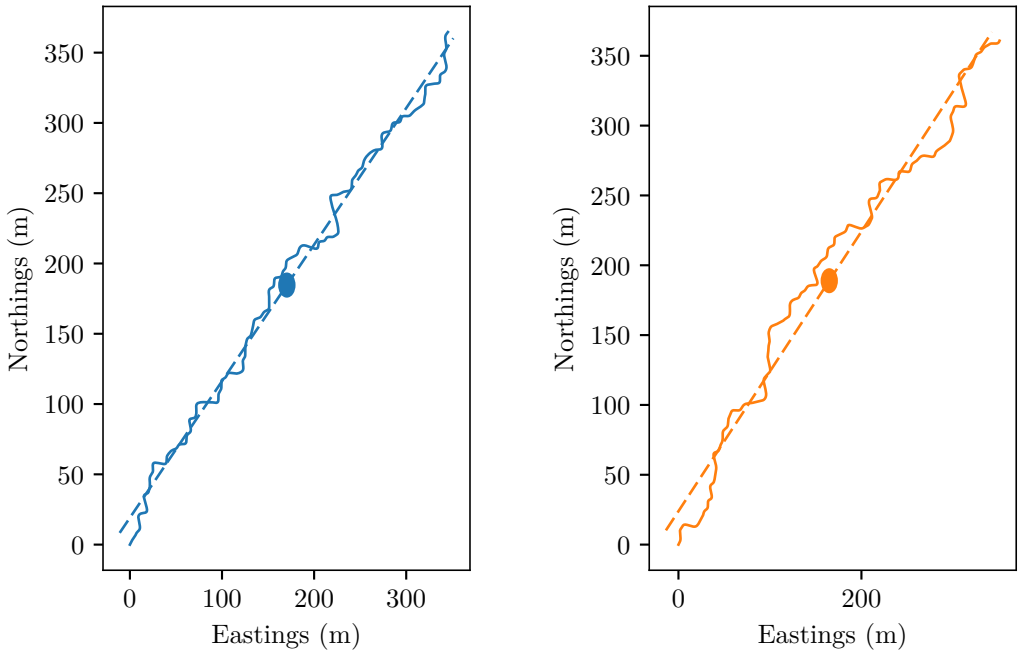
$$\hat{\mathbf{c}}_m, \hat{\mathbf{d}}_m = \underset{\mathbf{c}_m, \mathbf{d}_m}{\arg\min} \sum_{p=0}^{P_m - 1} \left| \mathbf{x}_m(p) - \left(p - \frac{P_m}{2} \right) \mathbf{d}_m - \mathbf{c}_m \right|^2, \tag{6.2}$$

where P_m is the total number of pings in the m^{th} track. The estimated straight tracks are then parameterised as

$$\hat{\mathbf{x}}_m(p) = \left(p - \frac{P_m}{2} \right) \hat{\mathbf{d}}_m + \hat{\mathbf{c}}_m. \tag{6.3}$$

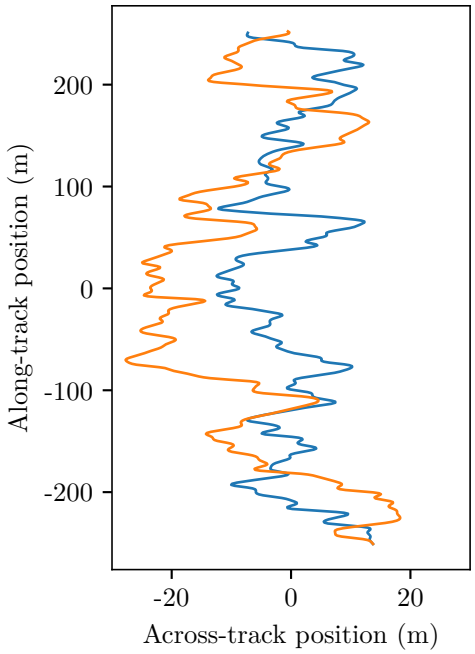
Two examples of this linear approximation are shown in Figures 6.1a and 6.1b. Each track is shown in its local coordinate system of metres east and north of a datum point. The linear approximation of each track given by (6.3) is shown as a dashed line, and the centre point is marked as a dot.

[1] These are sometimes — especially in radar literature — referred to as the azimuth and range directions respectively.



(a) Primary pass.

(b) Repeat pass.



(c) After transformation to the common coordinate system.

Figure 6.1: An example of the alignment of tracks to the common grid. Two tracks are shown in (a) and (b) with their local coordinate systems being the distance east and north of a datum point. The dashed lines show the linear approximation to the tracks, and the dots show the centre point. The centre point of the primary track is chosen as the origin of the common coordinate system, and the y or along-track axis is defined by the linear approximation to the primary track. The tracks after transformation into this coordinate system are shown in (c).

6.1.2 Coordinate system selection

One option for the coordinate system is to use the mean of the estimated linear tracks as the along-track axis. Having the chosen coordinate system in the 'centre' of the real tracks may be seen as advantageous. However, this becomes problematic when extra passes are added to the system after the initial processing: the coordinate system would change with the addition of another track, and each reconstructed image would have to be regenerated to fit the new coordinates. Additionally, it is common to treat the imagery from the primary pass as fixed, and adjust the subsequent (repeat) passes to align with it. In this context, it makes sense for the coordinate system to be based upon the primary pass.

Assume that the tracks exist in the same local coordinate system, i.e., if each track has a different datum or origin point, that the necessary corrections have been applied to account for this. Now let the estimated centre of the primary pass be the origin of the coordinate system, i.e.,

$$\mathbf{c} = \hat{\mathbf{c}}_0, \tag{6.4}$$

and the heading vector be the normalised version of its estimated displacement vector:

$$\begin{aligned} \mathbf{h} &= \frac{\hat{\mathbf{d}}_0}{\left|\hat{\mathbf{d}}_0\right|}, \\ &= \begin{bmatrix} h_x \\ h_y \\ h_z \end{bmatrix}. \end{aligned} \tag{6.5}$$

As discussed in Appendix A, a rotation can be represented as a 3×3 matrix; applying the rotation to a given coordinate vector is then a standard matrix multiplication. In other words, if $\mathbf{R}$ is a rotation matrix which rotates the heading vector $\mathbf{h}$ to align with the y-axis, then the original tracks can be transformed to the coordinate system by applying the calculated translation and rotation:

$$\mathbf{x}'_m(p) = \mathbf{R}(\mathbf{x}_m(p) - \mathbf{c}). \tag{6.6}$$

In the case where the sonar maintains the same depth throughout the track (i.e., $h_z = 0$), this realignment is a simple rotation around the z-axis. In this case, the heading angle of the sonar is given by

$$\theta_z = \arctan\left(\frac{h_x}{h_y}\right). \tag{6.7}$$

Taking the form of a rotation about the z-axis from (A.19), inserting θ_z as the rotation

angle, and applying trigonometric identities yields the rotation matrix

$$\mathbf{R} = \begin{bmatrix} h_y & -h_x & 0 \\ h_x & h_y & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{6.8}$$

In the more general case, an axis and angle of rotation need to be found. The axis is a vector perpendicular to both the initial and final vector. This can be calculated by using the cross product:

$$\begin{aligned} \mathbf{a} &= \mathbf{h} \times \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}^{\mathrm{T}}, \\ &= \begin{bmatrix} -h_z \\ 0 \\ h_x \end{bmatrix}. \end{aligned} \tag{6.9}$$

This can then be normalised to provide a unit axis of rotation

$$\hat{\mathbf{a}} = \begin{bmatrix} -\dfrac{h_z}{\sqrt{h_x^2 + h_z^2}} \\ 0 \\ \dfrac{h_x}{\sqrt{h_x^2 + h_z^2}} \end{bmatrix}. \tag{6.10}$$

Similarly, the angle of rotation can be found through the dot product as

$$\begin{aligned} \cos\theta &= \mathbf{h} \cdot \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}^{\mathrm{T}}, \\ &= h_y. \end{aligned} \tag{6.11}$$

Applying this axis and angle to the general form of (A.33) yields the rotation matrix

$$\mathbf{R} = \begin{bmatrix} h_y + \dfrac{h_z^2(1-h_y)}{h_x^2 + h_z^2} & -h_x & -\dfrac{h_x h_z(1-h_y)}{h_x^2 + h_z^2} \\ h_x & h_y & h_z \\ -\dfrac{h_x h_z(1-h_y)}{h_x^2 + h_z^2} & -h_z & \dfrac{h_x^2(1-h_y)}{h_x^2 + h_z^2} + h_y \end{bmatrix}. \tag{6.12}$$

Substituting $h_z = 0$ into this yields the simplified matrix of (6.8).

6.2 NAVIGATION ACCURACY

Ideally, the navigation data recorded by the platform carrying the sonar would be precise enough to register the reconstructed images to a sufficient accuracy for change detection. The majority of sonar systems are operated from an underwater towfish or

an AUV. From the length of the deployed cable and its direction as it leaves the ship, the position of a towfish can be estimated. This assumes that the cable run is straight and not affected by currents, the vessel's wake etc. Alternatively, an acoustic beacon on the towfish and a corresponding receiver on the vessel may be used to estimate the relative position of the two [Quinn 2014]. Any rotation of the towfish also alters the positions of the transmitter and receiver and must also be known or estimated. Techniques such as Kalman filtering [Kirlin et al. 1993] may be used to improve the estimates of the position and orientation of the towfish given a model of the system.

In the case of an AUV, global positioning system (GPS) — and similar technologies such as global navigation satellite system (GLONASS) or Galilieo — is not available underwater due to electromagnetic absorption[2]. Instead inertial navigation is used: starting from a known point (e.g., the launch point, or a GPS reading obtained on the surface), the velocity and orientation is measured and integrated to calculate the current position of the vehicle. Typically, this utilises a Doppler velocity log (DVL) and an inertial measurement unit (IMU) as the primary sensors [Kinsey, Eustice, and Whitcomb 2006]. Drift in the gyroscopes may be corrected by accelerometer and magnetometer readings [Kuch et al. 2012], with techniques such as Kalman filtering [Marins et al. 2001] or the Madgwick filter [Madgwick 2010] employed to improve accuracy. In regularly surveyed areas, acoustic transponders can be fixed to the seafloor around the area of interest. These form a long baseline (LBL) network which can be used to determine the position of the AUV [Pilbrow 2007]. Integrating these beacons with other sensors means fewer transponders are required [Willumsen, Hallingstad, and Jalving 2006].

Inertial systems are prone to drift. With good quality systems operating in ideal conditions the error in position is around 0.01 % of the distance travelled [Leonard et al. 1998] and can approach a 20 % drift error for low-grade systems [Munafò et al. 2014]. LBL positioning systems can achieve centimetre resolution in small areas [Leonard et al. 1998]. However, this is still too poor for accurate registration with high-resolution images.

6.3 THE MUD SONAR

Developed by the TNO,[3] the MUD SAS system is designed to operate in shallow, muddy estuarine and harbour environments [Vossen et al. 2012]. It has high frequency, low frequency and very low frequency transmitters paired with both horizontal and vertical arrays of receivers. It is unique in that it is mounted on the hull of a surface

[2]The carrier frequency of the GPS L1 band is 1575.42 MHz. At this frequency, the absorption coefficient of water is $\sim 1.2 \times 10^3\,\mathrm{m}^{-1}$, meaning over 90 % of the energy is absorbed in the first 2 mm and over 99.9 % in the first 6 mm.

[3]Nederlandse Organisatie voor Toegepast Natuurwetenschappelijk Onderzoek (Netherlands Organisation for Applied Scientific Research).

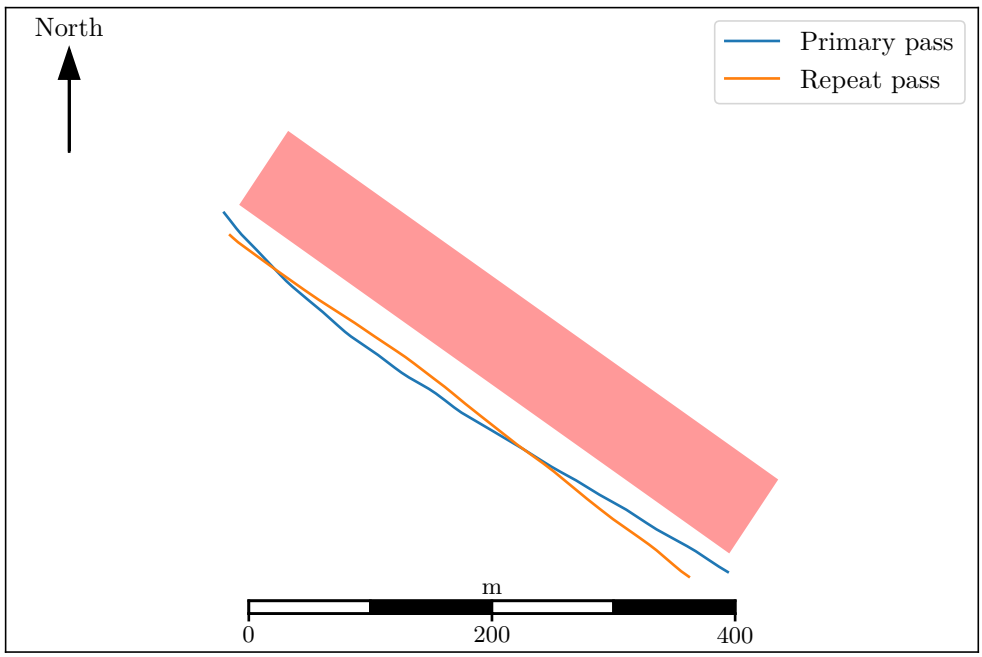


Figure 6.2: The track of the MUD sonar on the two passes analysed here. The shaded region shows the bounds of the reconstructed scene.

vessel meaning that its position relative to the ship is fixed. During the trials the data presented here were collected on, the ship was instrumented with an IXSEA PHINS inertial navigation system as well as real time kinematic (RTK) GPS.

RTK satellite navigation achieves high position accuracy by using the carrier wave of the underlying system (e.g., GPS). A base station at a precisely known location re-transmits the phase of the carrier signal it observes. The receiver units then compare this with the phase they observe. An accuracy of 1 cm ± 2 ppm horizontally and 2 cm ± 2 ppm vertically is achievable [Misra and Enge 2006], although the receiver must be within several tens of kilometres of the base station [Łapucha et al. 2011]. For surveys in coastal waters (such as the one presented here) this is not a major restriction.

Centimetre-level accuracy corresponds to about one-twentieth of the wavelength for the low frequency (6.5 kHz centre frequency and 5 kHz bandwidth) transmitter used in this study. This makes it a useful guide as to the registration accuracy achievable with high-quality position information. It is worth noting that the rotation (roll, yaw, pitch) of the ship will change the Cartesian offset between the sonar and the location of the GPS, and so there will be some degradation of the location accuracy due to uncertainties in the measured rotation angles.

Figure 6.2 shows the tracks followed by the MUD sonar on two runs over a portion of the Haringvliet inlet, near Rotterdam in the Netherlands. They are approximately 510 m (primary) and 470 m (secondary) long with an along-track sample spacing of

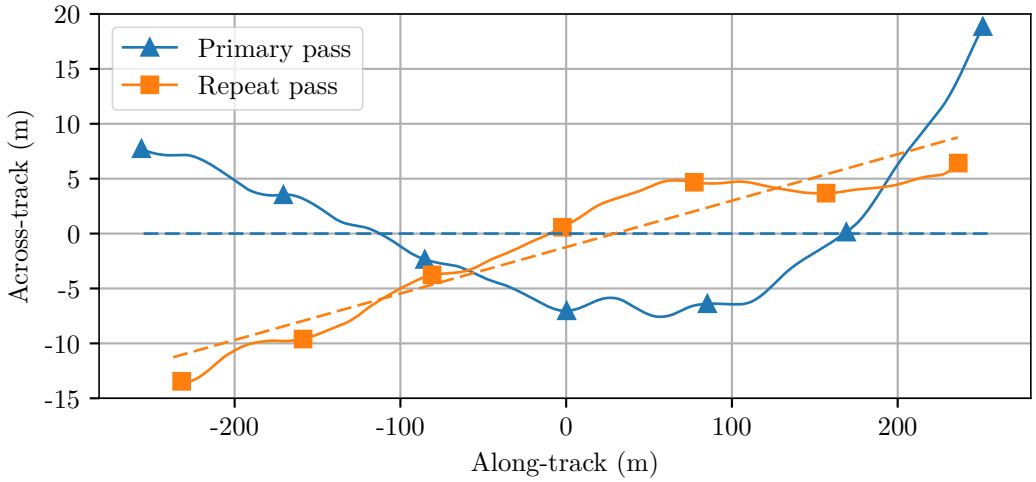


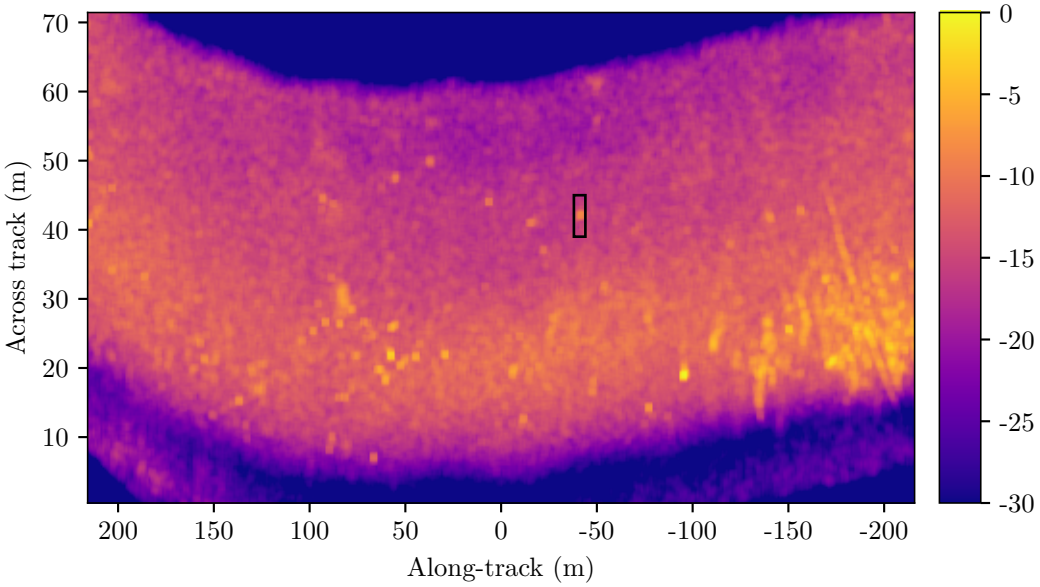
Figure 6.3: The tracks of the MUD sonar after the realignment procedure. Note that the axes have different scales which exaggerates the sway. The dashed lines show the linear approximation to the tracks; as defined, the approximation to the primary track now lies on the along-track axis. The angle between the two linear approximations (i.e., the difference in average heading between the passes) is 2.42°.

170 mm. This spacing is larger than the 150 mm wide projector, meaning that the synthetic aperture is undersampled (approximately $D/1$ sampling). However, the wide beamwidth and large bandwidth of the system smears the resulting grating lobes. The imaged scene, the reconstructed portion of which is marked by the shaded area in Figure 6.2, is a muddy seafloor at a depth of 13.6 m with a number of targets buried below the surface.

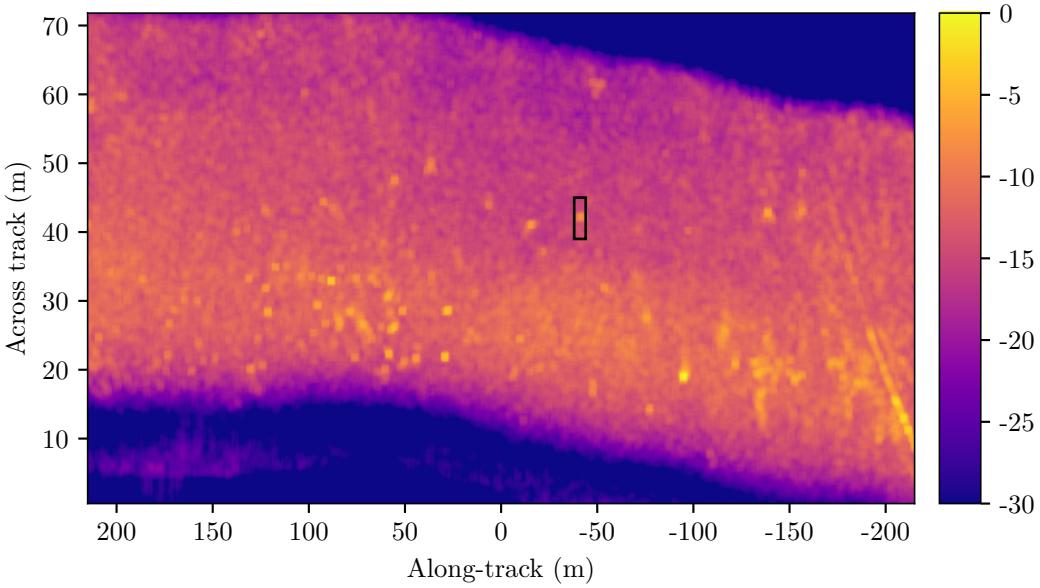
Performing the realignment procedure results in the tracks shown in Figure 6.3. Although the scale exaggerates it, there is a large sway in these tracks of up to 20 m from the mean path; this is a result of the sonar being hull mounted. Using these realigned tracks the data can be reconstructed onto a common grid. Figure 6.4 shows the results of using the backprojection algorithm. In order to reduce phase noise introduced by the linear interpolation used in the reconstruction, the raw data was first sinc interpolated by a factor of four [Barclay 2006]. Note that the images have been despeckled for display purposes by taking the RMS values over a 2 m by 2 m moving window. Subsequent processing and analysis is carried out without this despeckling.

The coherence between the two passes was calculated over a 2 m by 2 m sliding window. The magnitude of this coherence is shown in Figure 6.5; note that this image has been cropped to only consider the region where both passes of the sonar resulted in a high level of backscattered energy. The majority of this coherence is low, with only a few patches having a coherence greater than 0.5. The maximum observed coherence magnitude is 0.72.

The marked areas in Figure 6.4 were cropped out, and are shown in Figure 6.6.



(a) Primary pass.



(b) Repeat pass.

Figure 6.4: Reconstructed imagery from the two passes of the MUD sonar on the paths shown in Figure 6.2. The reconstruction was performed with the backprojection algorithm, and the 16 channels have been coherently summed. For display purposes despeckling was applied by taking the RMS values over a moving window. The colour scale is in decibels relative to the highest value in each image. The marked areas are further analysed in later figures.

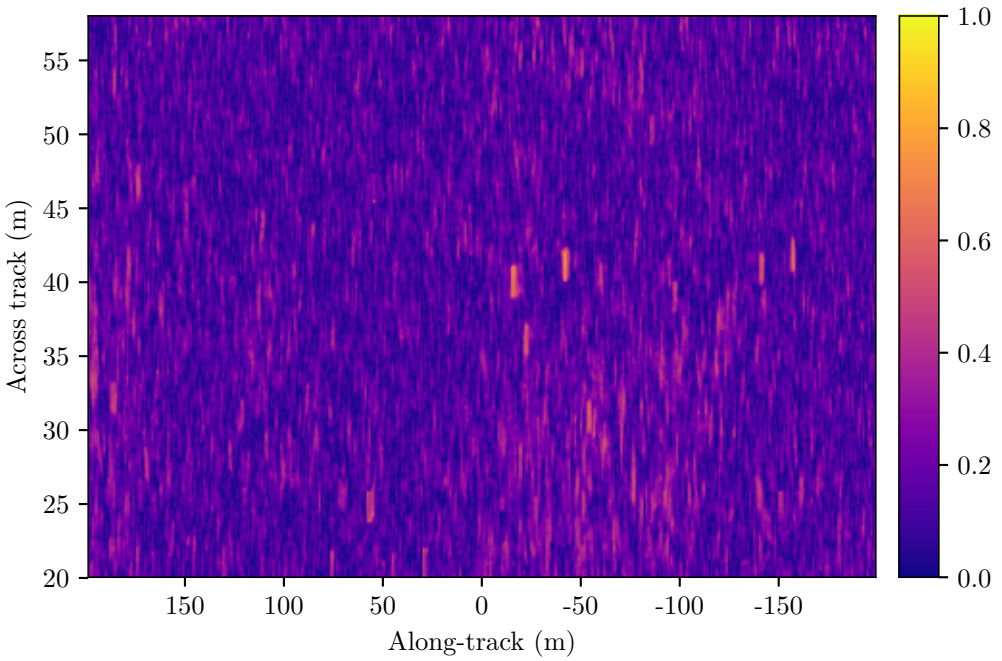
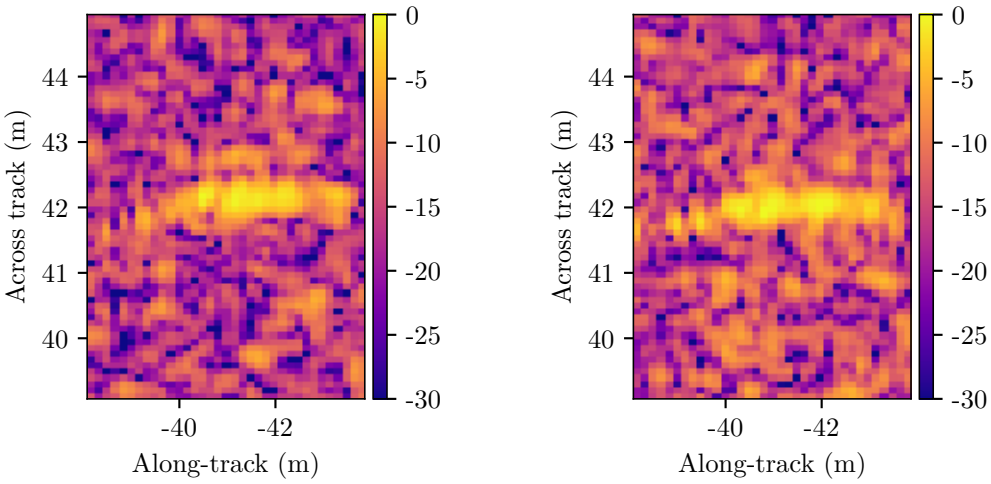


Figure 6.5: The magnitude of the repeat-pass coherence between the two passes of the MUD sonar shown in Figure 6.4. Note that this has been cropped to the region where both images contained a high level of backscattered energy.

This portion of the image consists of a smeared target amongst speckle noise. The area surrounding the target in the repeat pass image was correlated with the surrounding region of the primary pass image. The resulting correlation image is given in Figure 6.7. The peak of the correlation has a value of around 0.3, and is located at an across-track shift of 0.19 m and an along-track shift of 0.68 m. This shift shows that the misregistration error is high despite the use of the navigation data in the reconstruction process. The low peak value also suggests that there has been significant decorrelation between the two passes. A rotation error in the assumed track may cause this decorrelation (see Section 6.9.1), or it may be a temporal decorrelation between the passes. In either case, the large displacement between the two images shows that the navigation data is insufficiently accurate for registration of the images.

6.4 SYSTEM MODEL

Misregistration occurs when errors in the assumed track cause targets to be reconstructed in a different position. Note that an offset between the primary and secondary tracks does not cause misregistration. If both tracks are accurately known, and the reconstruction algorithm used to generate the imagery takes them into account, then common targets will be reconstructed into the same place in the images. As mentioned in Section 3.6.3, an offset will cause a loss of coherence due to baseline decorrelation



(a) Cropped area from primary pass.

(b) Cropped area from repeat pass.

Figure 6.6: The area of reconstructed MUD imagery marked in Figure 6.4. This consists of a speckle image with a smeared target running through the centre.

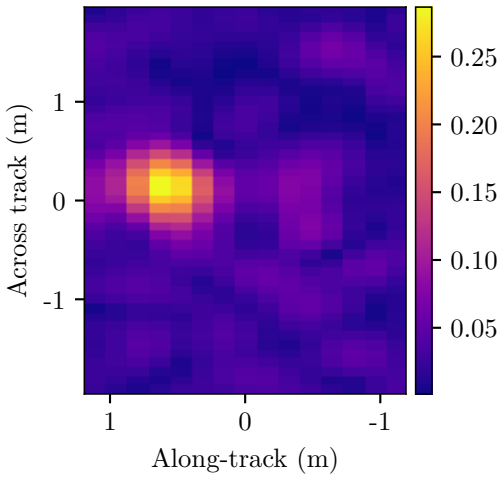
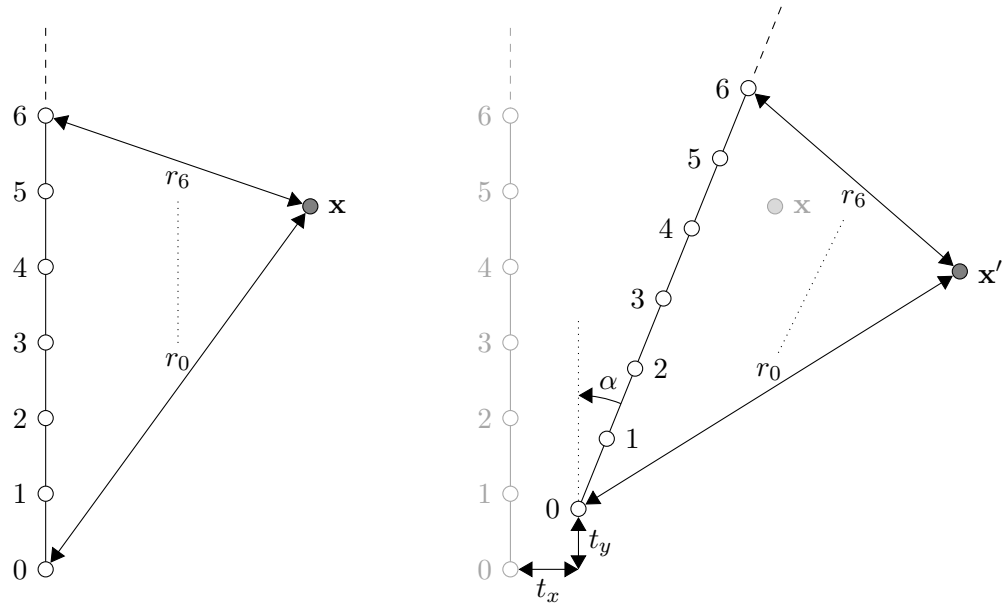


Figure 6.7: The correlation of the repeat-pass target shown in Figure 6.6b with the corresponding area in the primary image.



(a) With correct track. (b) With translation and heading errors.

Figure 6.8: The warping of the reconstructed image due to navigation errors. In (a), the correct track is used to reconstruct the image. Each ping is numbered, and the range from the ping positions to a particular location in the scene $\mathbf{x}$ is marked. As shown in (b), if the recorded navigation data has translation errors t_x and t_y and a rotation α from the true track, then the same point in the scene is reconstructed at position $\mathbf{x}'$ instead. Note that the direction of a positive heading angle error α is defined to match the right-hand rule.

and thus should be minimised.

Figure 6.8a shows a point in the scene $\mathbf{x}$ which has been insonified by a number of pings in the track, and thus the range from each of these pings to the point is known. This model assumes the tracks are reasonably straight so that any blurring in the reconstructed imagery is small. Using the correct track for the reconstruction procedure results in the point being placed at the correct position within the image. However, if the track has an across-track translational error t_x, an along-track translational error t_y, and a heading error α then, as illustrated in Figure 6.8b, the same point will appear at a different position $\mathbf{x}'$ in the reconstructed image. This warps the image, and, since the errors will be different for each pass over a scene, leads to misregistration between passes.

The displacement of a point in the reconstructed scene can be calculated from the model parameters. Let the point be represented by the vector

$$\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}. \tag{6.13}$$

The offset of the assumed secondary track from the actual track can also be represented by a vector,

$$\mathbf{t} = \begin{bmatrix} t_x \\ t_y \end{bmatrix}, \tag{6.14}$$

and the heading error can be used to form the rotation matrix

$$\mathbf{R}_\alpha = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}. \tag{6.15}$$

The warped position of the target due to these errors in the assumed track is then given by

$$\begin{aligned} \mathbf{x}' &= \begin{bmatrix} x' \\ y' \end{bmatrix}, \\ &= \mathbf{R}_\alpha \mathbf{x} + \mathbf{t}, \end{aligned} \tag{6.16}$$

and the error in the position is thus

$$\begin{aligned} \Delta\mathbf{x} &= \mathbf{x} - \mathbf{x}', \\ &= \mathbf{x} - \mathbf{R}_\alpha \mathbf{x} - \mathbf{t}. \end{aligned} \tag{6.17}$$

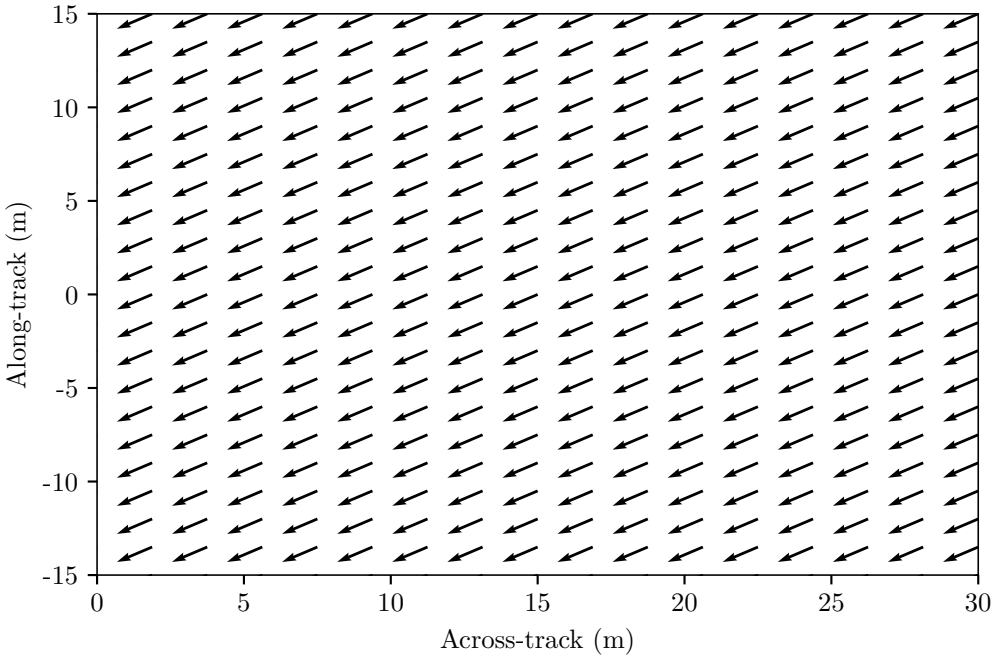
Breaking this into its across- and along-track components and expanding gives

$$\begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = \begin{bmatrix} x - x\cos\alpha + y\sin\alpha - t_x \\ y - x\sin\alpha - y\cos\alpha - t_y \end{bmatrix}. \tag{6.18}$$

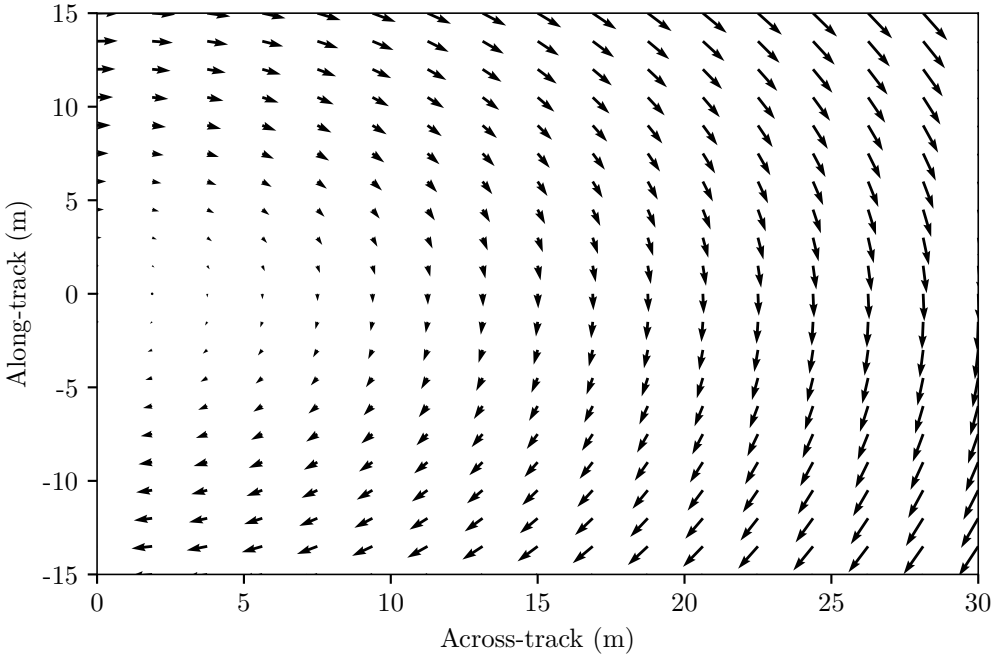
This error can be visualised as a *displacement field*, a set of vectors showing the error at different points in the image. Figure 6.9a shows the displacement field for a pure translation error, in this case $t_x = 1.2\,\text{m}$ and $t_y = 0.8\,\text{m}$. The tail of each arrow sits at the misregistered or warped point $\mathbf{x}'$ and the tip or head sits at the true ($\mathbf{x}$) point, i.e., the arrows show how the image would need to be corrected to be registered with the primary pass. An example of the displacement field for a pure rotation of $\alpha = 3°$ is given in Figure 6.9b. As might be expected, the further away from the sonar the larger the displacement due to the rotation is. Two examples of combined translation and rotation errors are presented in Figure 6.10. The first example (Figure 6.10a) combines the errors from the two previous examples, i.e., $t_x = 1.2\,\text{m}$, $t_y = 0.8\,\text{m}$, and $\alpha = 3°$, and shows that the resulting displacement field is the linear sum of the two individual displacement fields. This is expected from (6.18).

6.5 CORRELATION PEAKS

In order to calculate the track error parameters, some estimate of the displacement field needs to be generated from the two images. The method proposed here is the

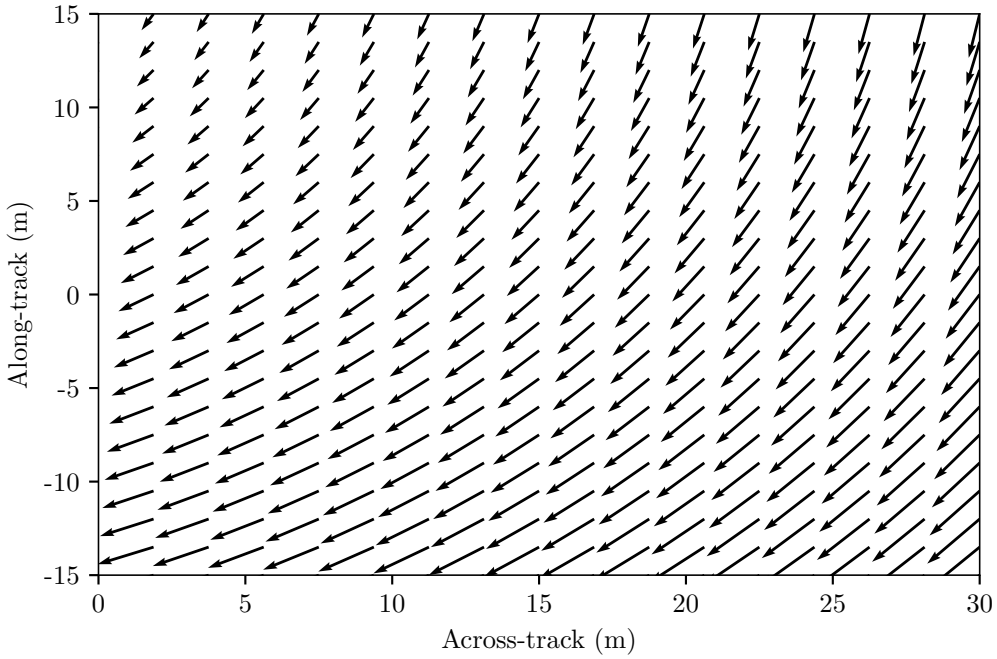


(a) Pure translation: $t_x = 1.2\,\text{m}$, $t_y = 0.8\,\text{m}$, $\alpha = 0°$.

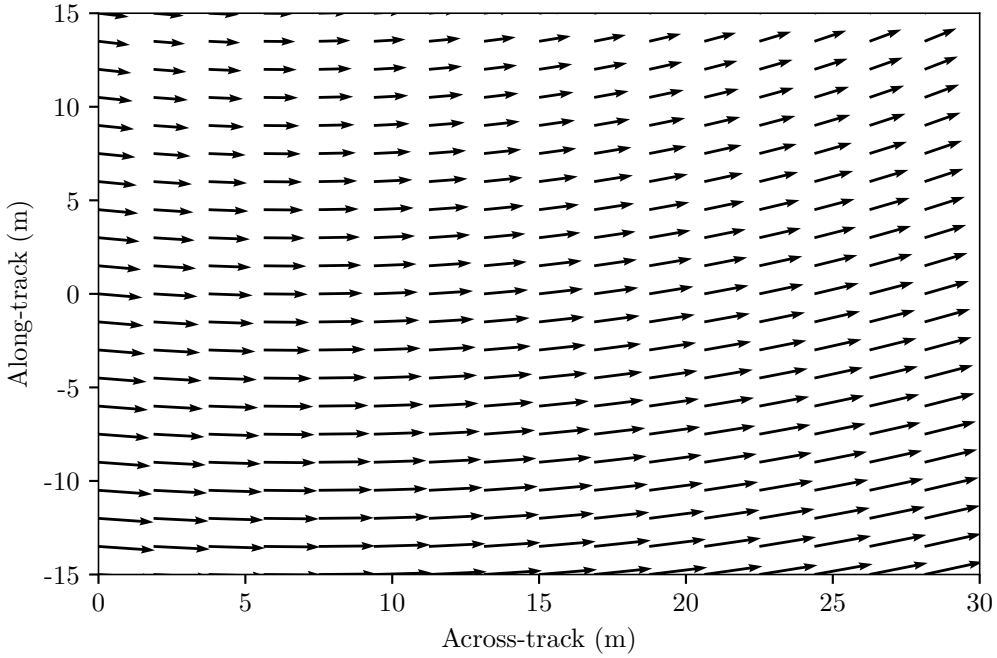


(b) Pure rotation: $t_x = 0\,\text{m}$, $t_y = 0\,\text{m}$, $\alpha = 3°$.

Figure 6.9: Examples of the displacement field for (a) translation and (b) rotation errors. The tail of the arrows are at the misregistered ($\mathbf{x}'$) positions of a point in the scene, and the tips are at the correct ($\mathbf{x}$) positions of the same point. In other words, the arrows show how the image needs to be corrected to be registered with the image from the primary pass.



(a) $t_x = 1.2\,\mathrm{m}$, $t_y = 0.8\,\mathrm{m}$, $\alpha = 3°$.



(b) $t_x = -1.5\,\mathrm{m}$, $t_y = 0.2\,\mathrm{m}$, $\alpha = -1.8°$.

Figure 6.10: Examples of the displacement field when both translation and rotation errors occur in the same track. In (a), the errors are the combination of those used in Figures 6.9a and 6.9b and the displacement field is a linear combination of those two fields.

block correlation of the images. The repeat image is divided into a number of adjacent blocks. For each repeat-pass block, a larger concentric block is extracted from the primary image. These two blocks are then correlated; the difference in size between the blocks defines the distance over which this correlation takes place. For example, a 2 m by 2 m repeat pass block correlated with a concentric 8 m by 8 m primary pass block will result in a correlation image over a ± 3 m search space in each direction from the centre of the repeat pass block. Locating the position of the peak of this correlation image gives an estimate of the displacement field at this location, while the value of the peak provides a metric as to the strength of the match.

6.6 ESTIMATION OF MODEL PARAMETERS

Given a set of observed displacements, the problem is therefore to solve (6.18) for the corresponding translation and heading errors in the track. The observations will not be perfect; any source of decorrelation — be it due to the scene changing or noise in the measurements — will cause errors. It will also be noted that this forms an overdetermined problem. There are $M \times N$ observations of the correlation peaks, one from each block, and only three parameters t_x, t_y, and α to be estimated. Two techniques are presented in this section, a linear least squares approach for small heading errors in Section 6.6.1 and an optimisation problem for larger heading errors in Section 6.6.2.

6.6.1 Small heading error

If the heading error α is small, then (6.18) can be linearised using the small-angle approximations $\sin\alpha \approx \alpha$ and $\cos\alpha \approx 1$. This gives the misregistration due to the track errors as

$$\begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \approx \begin{bmatrix} y\alpha - t_x \\ -x\alpha - t_y \end{bmatrix}. \tag{6.19}$$

The error parameters can be used to form a vector

$$\mathbf{e} = \begin{bmatrix} t_x \\ t_y \\ \alpha \end{bmatrix}. \tag{6.20}$$

If N observations of the misregistration are available, then an observation vector can be created:

$$\mathbf{b} = \begin{bmatrix} \Delta\widetilde{x}_1 \\ \Delta\widetilde{y}_1 \\ \Delta\widetilde{x}_2 \\ \Delta\widetilde{y}_2 \\ \vdots \\ \Delta\widetilde{x}_N \\ \Delta\widetilde{y}_N \end{bmatrix}, \tag{6.21}$$

where $\Delta\widetilde{x}_n$ and $\Delta\widetilde{y}_n$ are the n^{th} observations of the across- and along-track misregistrations respectively. Using (6.19), a linear least squares system

$$\mathbf{A}\mathbf{e} = \mathbf{b} \tag{6.22}$$

can then be formulated, where

$$\mathbf{A} = \begin{bmatrix} -1 & 0 & y_1 \\ 0 & -1 & -x_1 \\ -1 & 0 & y_2 \\ 0 & -1 & -x_2 \\ \vdots & \vdots & \vdots \\ -1 & 0 & y_N \\ 0 & -1 & -x_N \end{bmatrix} \tag{6.23}$$

and x_n and y_n are the correct across- and along-track locations of the observation points respectively. Rearranging (6.22) yields an estimate for $\mathbf{e}$:

$$\widehat{\mathbf{e}} = \mathbf{A}^{+}\mathbf{b}, \tag{6.24}$$

where $\mathbf{A}^{+} = \left(\mathbf{A}^{\mathrm{T}}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}$ is the Moore-Penrose pseudoinverse [Moore 1920; Penrose 1955] of $\mathbf{A}$. This estimate will be less accurate if the observations have errors in them. If some measure of the accuracy of each observation is available, then a diagonal weighting matrix can be formed:

$$\mathbf{W} = \begin{bmatrix} w_1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & w_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & w_2 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & w_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & w_N & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & w_N \end{bmatrix}, \tag{6.25}$$

where w_n is the weighting of the n^{th} observation. For example, misregistration measurements made via correlation could use the magnitude of the peak correlation of each observation as the weighting. Using the weighting matrix, (6.22) can be modified to give a weighted linear least squares system,

$$\mathbf{WAe} = \mathbf{b}. \tag{6.26}$$

The estimate of the parameters is then

$$\widehat{\mathbf{e}} = \mathbf{A}_{\mathbf{w}}^{+}\mathbf{b}_{\mathbf{w}}, \tag{6.27}$$

where $\mathbf{A}_{\mathbf{w}}^{+}$ is the pseudeoinverse of $\mathbf{WA}$ and $\mathbf{b}_{\mathbf{w}} = \mathbf{Wb}$.

6.6.2 Large heading error

As the heading error α increases, the small angle approximation becomes invalid. This means that the model (6.18) can no longer be linearised. Instead, an optimisation problem can be posed to find estimates of the track error parameters:

$$\hat{t}_x, \hat{t}_y, \hat{\alpha} = \underset{t_x, t_y, \alpha}{\arg\min}\, \epsilon(t_x, t_y, \alpha), \tag{6.28}$$

where $\epsilon(t_x, t_y, \alpha)$ is some measure of the error between the observed misregistrations and the model predictions for the given parameters. One potential error function is the total squared error. The squared error for the n^{th} observation is given by

$$\epsilon_n(t_x, t_y, \alpha) = (\Delta x_n - \Delta\widetilde{x}_n)^2 + (\Delta y_n - \Delta\widetilde{y}_n)^2, \tag{6.29}$$

where Δx_n and Δy_n are the predicted misregistrations from (6.18), and $\Delta\widetilde{x}_n$ and $\Delta\widetilde{y}_n$ are the observed values. The overall squared error is then

$$\epsilon(t_x, t_y, \alpha) = \sum_{n=1}^{N} \epsilon_n(t_x, t_y, \alpha). \tag{6.30}$$

As with the previous section, errors introduced by lower-accuracy observations can be mitigated by assigned a weighting to each observation. This results in (6.30) being modified to

$$\epsilon(t_x, t_y, \alpha) = \sum_{n=1}^{N} w_n \epsilon_n(t_x, t_y, \alpha). \tag{6.31}$$

6.7 SIMULATED SONAR PARAMETERS

In order to test the performance of the presented registration algorithms, a number of repeat-pass datasets were simulated. The parameters used for these simulations are

Parameter	Value
Transducer	
Width	40 mm
Height	20 mm
Tilt	12° down
Transmitted signal	
Type	Linear FM up-chirp
Duration	3.3 ms
Centre frequency	120 kHz
Bandwidth	30 kHz
Window	None
Wavelength	11.1 mm – 14.3 mm
Receiver	
Ping spacing	10 mm
Sample rate	31.25 kHz
Samples per ping	1,658

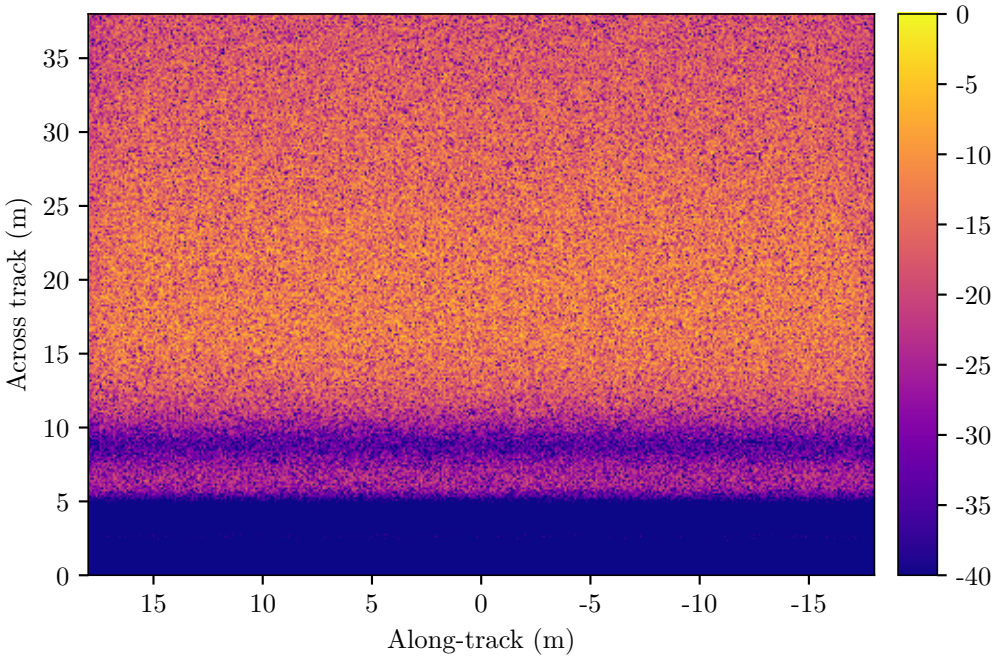
Table 6.1: Parameters of the sonar used to generate simulated data for testing the registration algorithms presented in this chapter. The same transducer was used for both transmitting and receiving.

given in Table 6.1. To reduce the effects of grating lobes on the repeat-pass coherence, the ping spacing was chosen to be one-quarter of the along-track width of the transducer, i.e., $D/4$ sampling was used. The size of the resolution cell in the across- and along-track directions can be calculated from (1.1) and (1.5) respectively as $\delta_r = 25\,\text{mm}$ and $\delta_y = 20\,\text{mm}$. From this, the desired registration accuracy to minimise coherence loss (10 % of the resolution cell) is 2.5 mm in the across-track direction and 2 mm in the along-track direction.

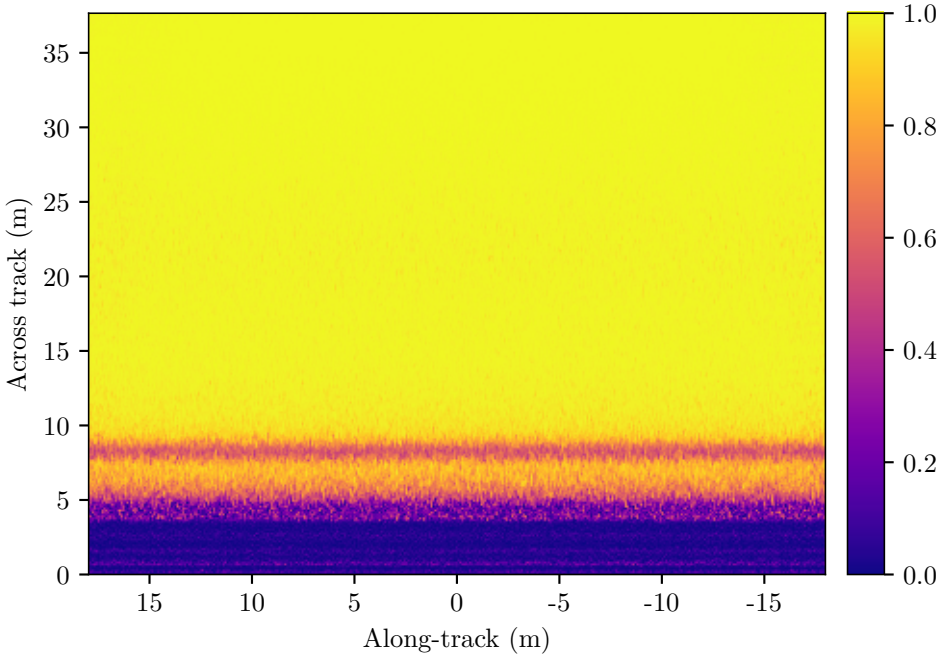
6.8 TRANSLATION ERRORS

A bland scene was simulated for two passes of the experimental sonar. The repeat pass was offset by -1 m (i.e., one metre further away from the scene) from the primary pass. The reconstructed image for the primary pass is shown in Figure 6.11a. The low energy area for across-track positions below 10 m is due to this area of the scene not being in the main lobe of the sonar's vertical beampattern. When the repeat pass is reconstructed with the correct track and no additive noise, the coherence between the images is dominated by the baseline decorrelation. As shown in Figure 6.11b, the small separation between the tracks results in minimal degradation of the coherence.

For peak detection, the standard deviation of the errors in the estimated location of the peak is inversely proportional to its bandwidth, i.e., a higher bandwidth will yield a more accurate estimate [Quazi 1981]. Equivalently, the errors are proportional to the width of the peak (so a narrower peak will result in a lower standard devia-



(a) The magnitude of the primary pass. The colour scale is in decibels relative to the highest magnitude.



(b) The degree of coherence between the primary and repeat passes when the correct track is used to reconstruct the repeat pass.

Figure 6.11: The (a) magnitude of the primary pass and (b) repeat-pass coherence when the correct tracks are used to reconstruct the imagery. The bands of low energy and coherence respectively are due to those areas of the scene being outside the main lobe of the sonar beampattern.

tion). When detecting peaks in the correlation of two sonar images, the shape of the peak is determined by the autocorrelation of the system impulse response. For SAS speckle imagery, the across-track autocorrelation is a function of the bandwidth of the transmitted signal, and the along-track autocorrelation is a function of the width and shape of the transmitting and receiving apertures [Fortune 2005]. The autocorrelation of the impulse response of the experimental sonar used here is shown in Figure 6.12. Calculating between the points where the autocorrelation is 0.5, the across-track autocorrelation has a peak width of 84 mm and the along-track autocorrelation has a peak width of 45 mm. The errors in the across-track direction are therefore expected to be twice the errors in the along-track direction.

6.8.1 Generating translation errors

In order to test the performance of the estimators in the presence of translation errors, a set of 343 repeat-pass images with translation errors was created. For each image, an error between -3 m and 3 m in each direction was randomly chosen from a uniform distribution. The correct track was then shifted by the selected distance. The data was backprojected using this incorrect track to give an image misregistered by the given amount. Figure 6.13 shows the values of these errors.

6.8.2 Quadratic interpolation

Each of the repeat-pass images was divided in 2 m by 2 m blocks, and each of these blocks was correlated with a concentric 8 m by 8 m block in the primary image. This means that an area 3 m by 3 m was searched for each block.[4] The position of the largest correlation magnitude was found, and then a quadratic surface was fitted to the eight pixels surrounding it. The peak position of this quadratic surface was then taken to be the position of the largest magnitude. This is functionally identical to performing a quadratic interpolation on the data and finding the largest magnitude of the interpolated data. An example of the detected peaks for a repeat-pass image with $t_x = 0.802\,\text{m}$ and $t_y = -0.549\,\text{m}$ is shown in Figure 6.14. It will be observed that the arrows indicating the position of the maximum correlation magnitude point are in the opposite direction to t_x and t_y. This is because the correlation peak indicates the distance the image needs to be shifted to correct the misregistration, i.e., to undo the effects of t_x and t_y.

The errors of the peak positions found by this correlation and quadratic interpolation approach are shown in Figure 6.15, and the corresponding standard deviations are given in Figure 6.16. For the majority of the test cases, the error is below 0.5 mm

[4] The repeat-pass block can only be shifted by 3 m in any direction before its edges start leaving the primary image block.

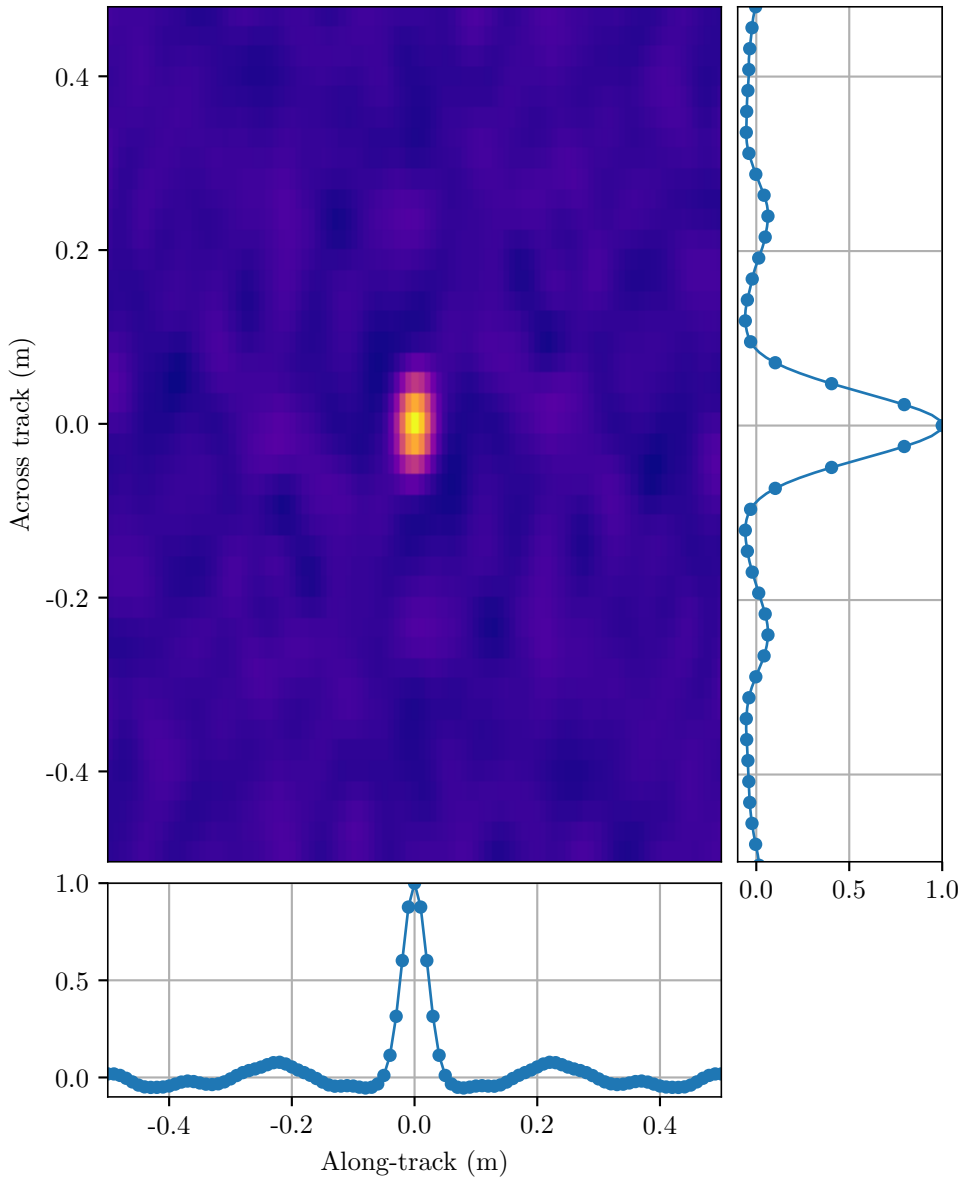


Figure 6.12: The autocorrelation of the impulse response of the test system. The image shows the autocorrelation over a 1 m by 1 m area while the two plots show slices through the centre of the autocorrelation in the corresponding directions. The widths of the peaks (measured between the points where the autocorrelation is 0.5) are 84 mm in the across-track direction and 45 mm in the along-track direction.

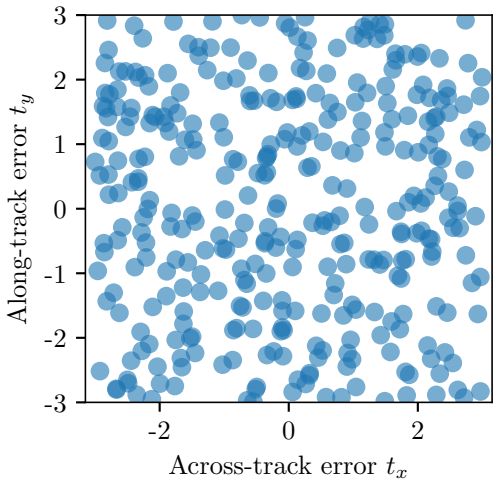


Figure 6.13: The set of 343 translation errors used for testing the registration correction procedure. The along-track and across-track errors were independently chosen from a uniform distribution between −3 m and 3 m.

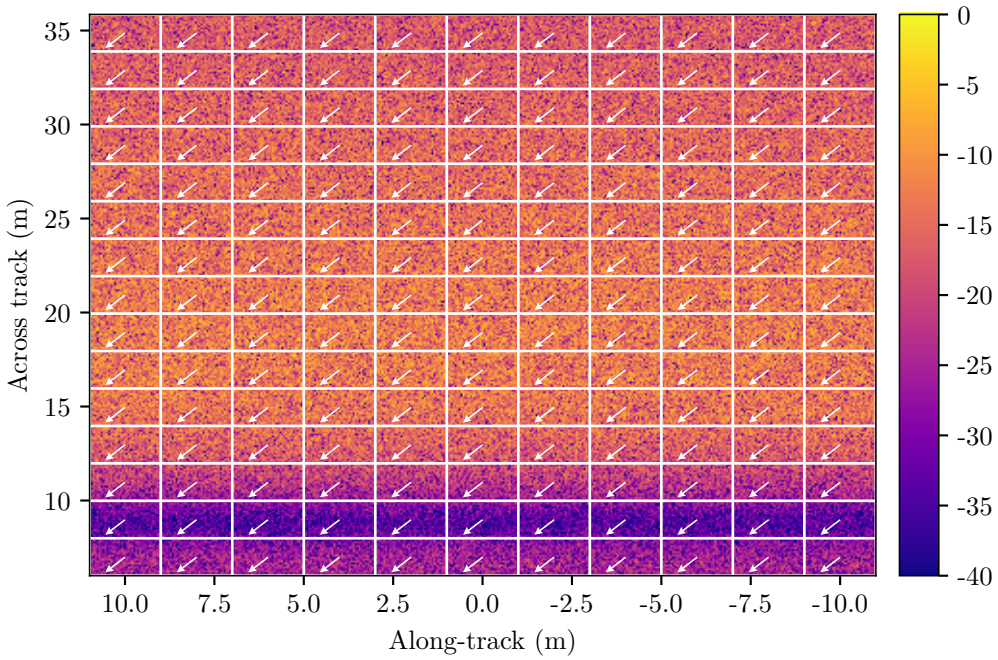
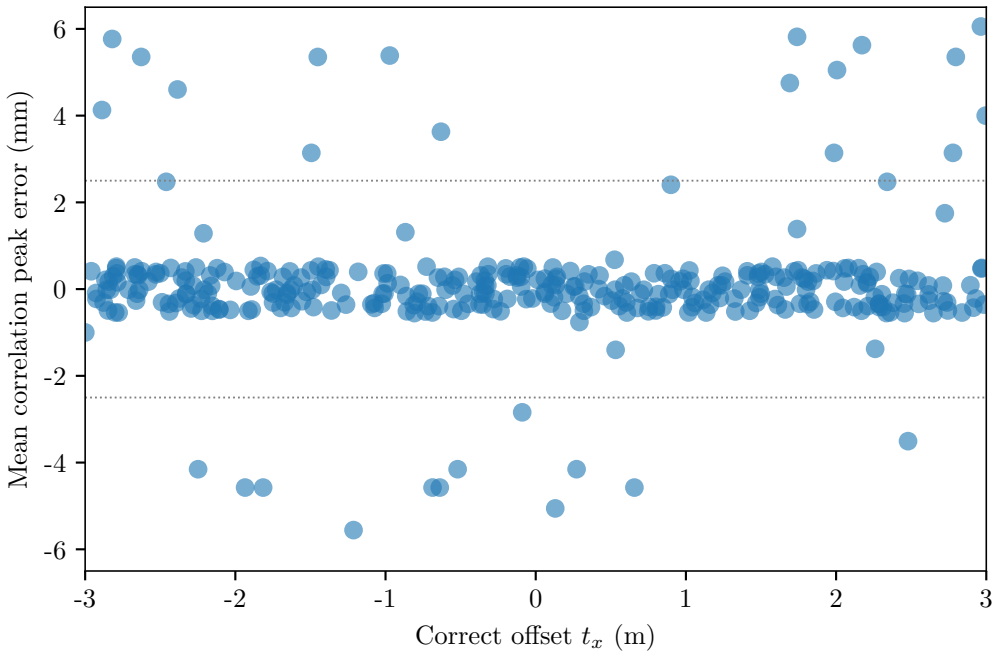


Figure 6.14: An example of peak detection from a repeat-pass image with a translation error. The underlying image is the image magnitude in decibels relative to the largest value. The tail of the arrows is at the centre of the blocks, and the tip shows the shift for which the maximum correlation was found. In this case the across-track error $t_x = 0.802\,\text{m}$ and the along-track error $t_y = -0.549\,\text{m}$. Note that the shifts point in the opposite direction to these values: if the image has shifted one way, correcting it requires shifting in the opposite direction.

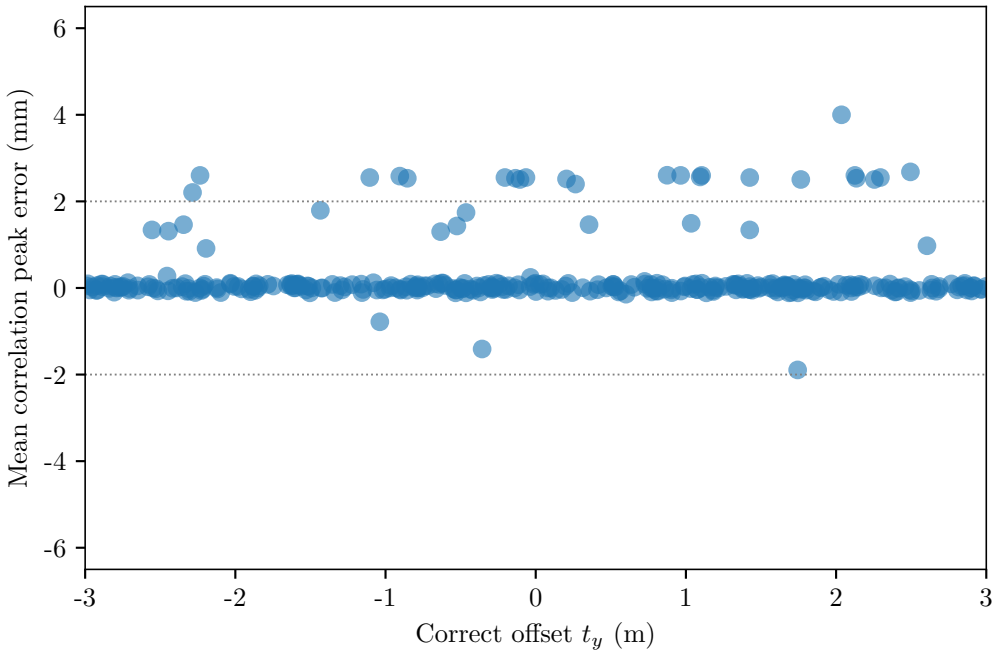
in the across-track case and 0.25 mm in the along-track case. These errors are approximately 0.02 and 0.0125 of a resolution cell respectively. The standard deviations are similarly proportioned: the majority are below 1 mm in across-track and 0.5 mm in along-track. This fits with the expectation that, based upon the autocorrelation of the system impulse response, the across-track errors would be twice the along-track errors.

The least squares and weighed least squares estimators were applied to the displacements found in each of the trials. The resulting errors in the estimated parameters are presented in Figure 6.17. The least squares errors are identical to the errors in the peak location estimates; this is not surprising as, for a purely translational case, the mean is the minimum error solution. Weighted least squares only reduced the errors by less than 10 %. The optimisation and weighted optimisation approaches gave identical results to their corresponding least-squares estimator. This is expected as with a rotation error $\alpha = 0$, the linearised model (6.19) is an exact fit to the full model.

The maximum errors are around 12 times larger than the majority. In the across-track case, the maximum error corresponds to 0.25 of the resolution cell size, and in the along-track direction it is 0.15 of the resolution cell size. Histograms of the errors in the detected across-track peak location are shown in Figure 6.18a for the worst-case ($t_x = 2.964$ m, mean error of 6.05 mm) and best-case ($t_x = 0.42$ m, mean error of less than 1 μm) estimates. The corresponding values of the peak correlation are given in Figure 6.18b. In the best-case scenario the errors are clustered around zero, while for the worst-case scenario approximately half of the errors are around zero and the other half are at an error of -12 mm, or half of the across-track image resolution. Although simple and efficient to implement, quadratic interpolation is biased, with the worst variance of the errors occurring at half the original sample spacing [Boucher and Hassab 1981]. This leads to the hypothesis that these large errors occur when the translation falls halfway between two samples and so the quadratic interpolator is unable to estimate the position accurately. This is exaggerated by the fact that since all the blocks have the same translation in this test, the estimate error occurs in a significant portion of the blocks and thus biases the overall estimate.

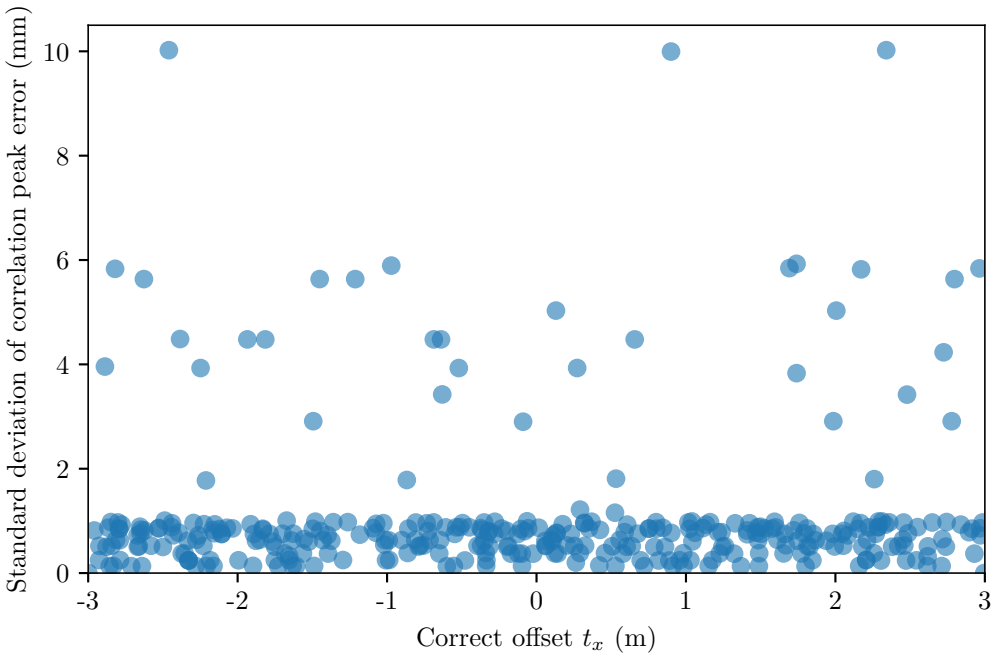


(a) Across-track error.

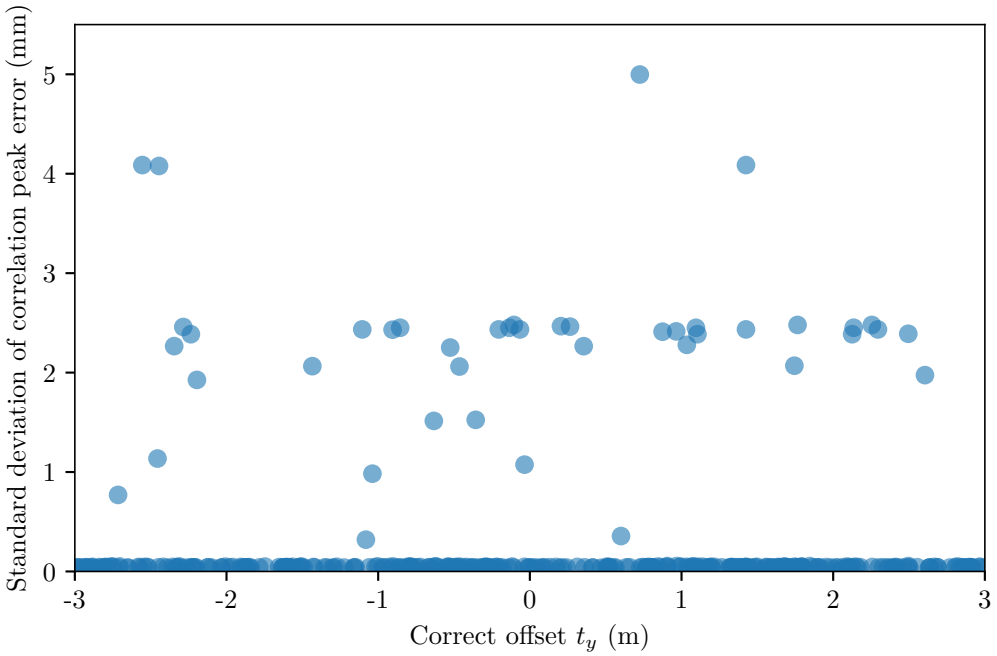


(b) Along-track error.

Figure 6.15: The mean error when estimating the position of the correlation peaks with quadratic interpolation. In both plots the x-axis shows the actual offset that was used in the trials. Each trial involved 165 blocks; the y-axes show the mean of the errors between the estimated correlation peak positions and the actual offsets over these blocks for each trial. The target maximum misregistration of 10 % of the sonar resolution cell is marked by the dotted lines.

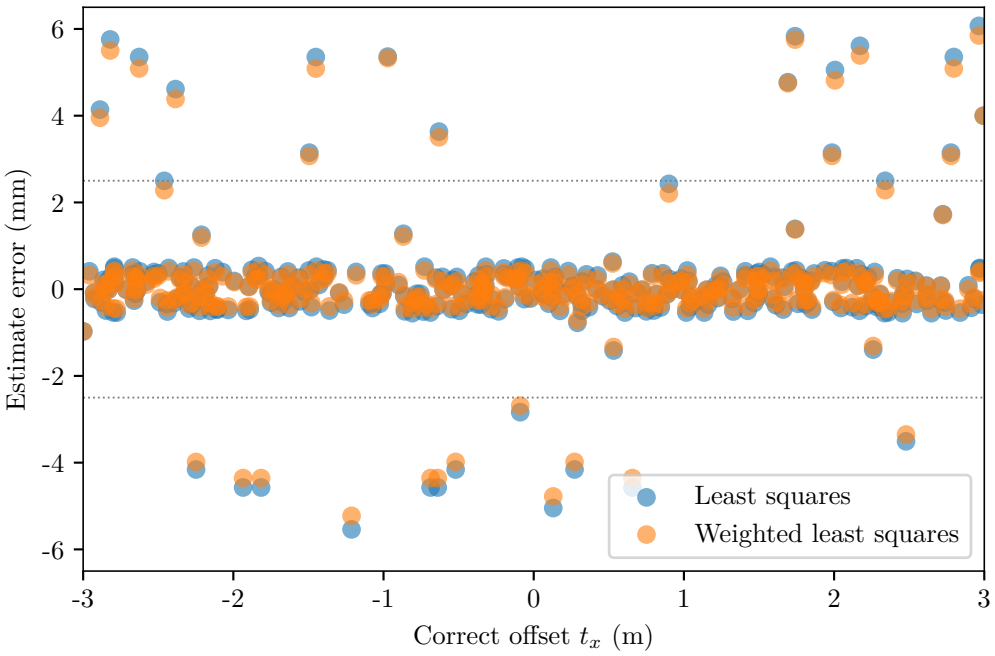


(a) Across-track error.

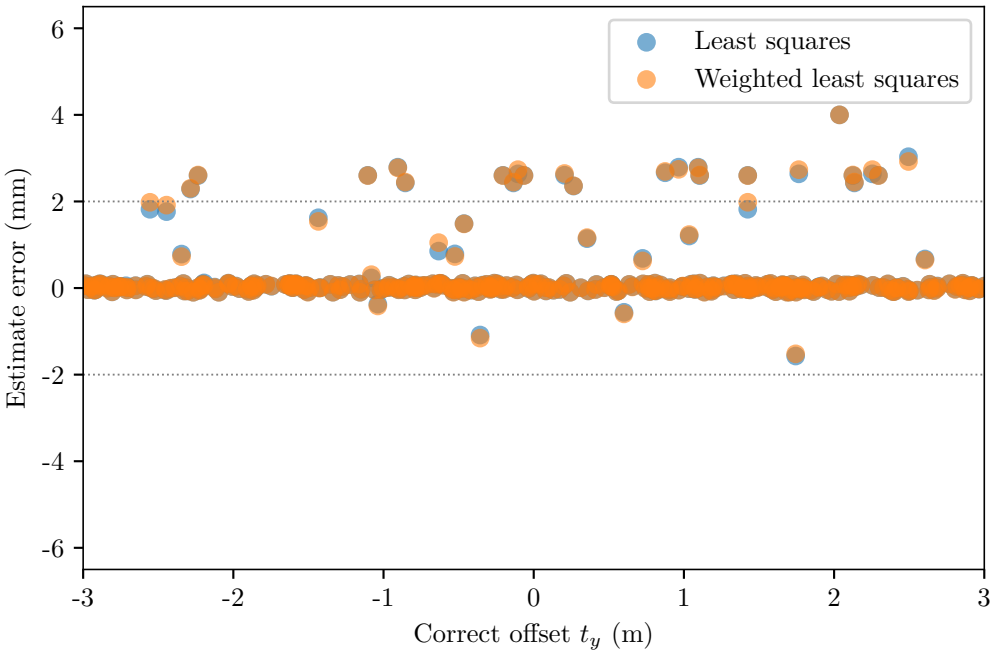


(b) Along-track error.

Figure 6.16: The standard deviation of the error when estimating the position of the correlation peaks with quadratic interpolation. In both plots the x-axis shows the actual offset that was used in the trials. Each trial involved 165 blocks; the y-axes show the standard deviation of the errors between the estimated correlation peak positions and the actual offsets over these blocks for each trial.



(a) Across-track error.



(b) Along-track error.

Figure 6.17: The errors in the translation parameter estimates when using quadratic interpolation to find the peak correlation position. Results for both the least squares and weighted least squares estimators are shown here; the optimisation and weighted optimisation estimates are identical to these. The target maximum misregistration of 10 % of the sonar resolution cell is marked by the dotted lines.

Considering the peak correlation values of Figure 6.18b, there are a number of blocks with a high correlation (greater than 0.9) and a number with a lower correlation (between 0.65 and 0.80). If these lower values correspond to blocks with a large error in the estimated position, then thresholding the positions to be used in the parameter estimation step may improve the estimate. The results of this thresholding are given in Figure 6.19. The errors are reduced compared to using all the detected displacements, but there are still a number of outlying cases.

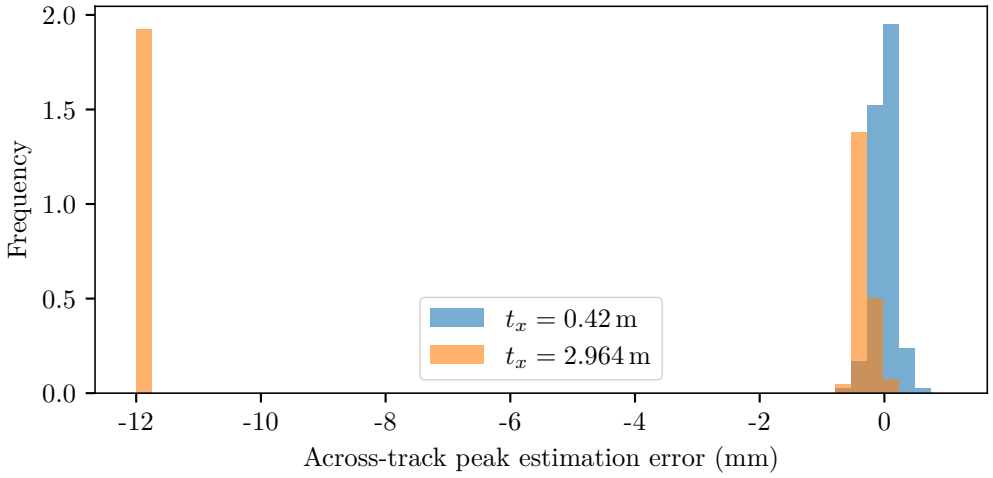
6.8.3 Sinc interpolation

Another method of estimating the position of the peak is to use a sinc interpolator to oversample the correlation image. The reduced sample spacing results in the maximum value in the image being a closer approximate to the true peak of the correlation. It is also possible to apply the quadratic interpolator after first performing sinc interpolation. Although the bias of the quadratic interpolator will still result in errors in the detected peak positions, they will be smaller than with quadratic interpolation alone due to the initial oversampling.

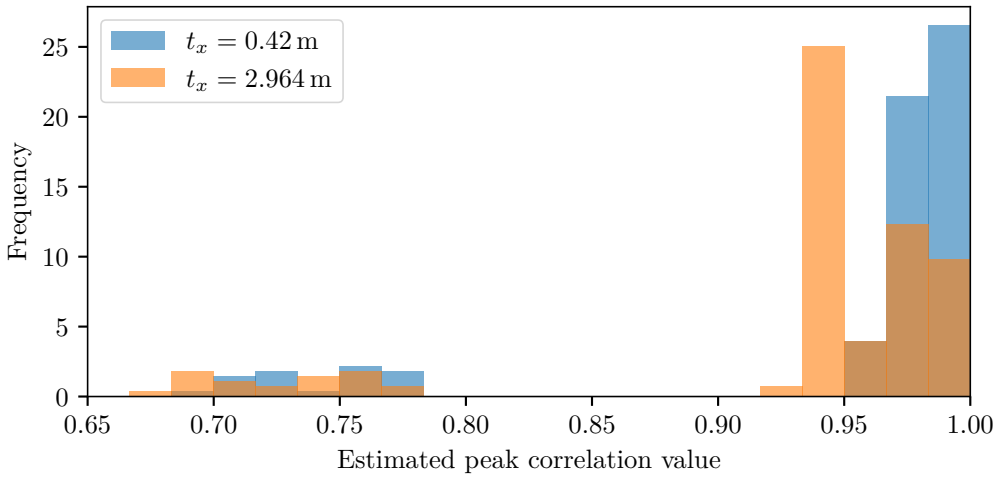
These approaches were tested for both $2\times$ and $4\times$ sinc interpolation. The resulting errors in the parameter estimation — which was performed using the least-squares estimator — are shown in Figures 6.20 and 6.21 respectively. With two-times sinc oversampling, the average error is noticeably worse than with the original quadratic interpolation, and the maximum error is only marginally lower. Applying the quadratic interpolator to the oversampled image leads to the average error being similar to solely using quadratic interpolation, but with a halving of the maximum error.

In the case of four-times sinc oversampling, the range of errors is reduced compared to using two-times oversampling. However, the average error is still worse than with just quadratic interpolation. Again, applying the quadratic interpolator to the oversampled image reduces the average error to a similar level to sole quadratic interpolation, and the maximum error is reduced further compared to both the quadratic-only and two-times sinc oversampling plus quadratic interpolation cases.

In theory, the oversampling factor could continue to be increased to obtain a further reduction in errors. However, this would follow the law of diminishing returns: at some point, the added costs in terms of computational and memory requirements would outweigh the reduction in errors. There is no appreciable difference in the average errors of the two- and four-times sinc plus quadratic cases shown. The main benefit of the sinc interpolation is the lowering of the maximum error in the estimates of the parameters. The maximum error with two-times sinc oversampling followed by quadratic interpolation is approximately equal to the desired 0.1 of a resolution cell threshold. As such, this method of peak location will be used for the remainder of the results in this chapter.

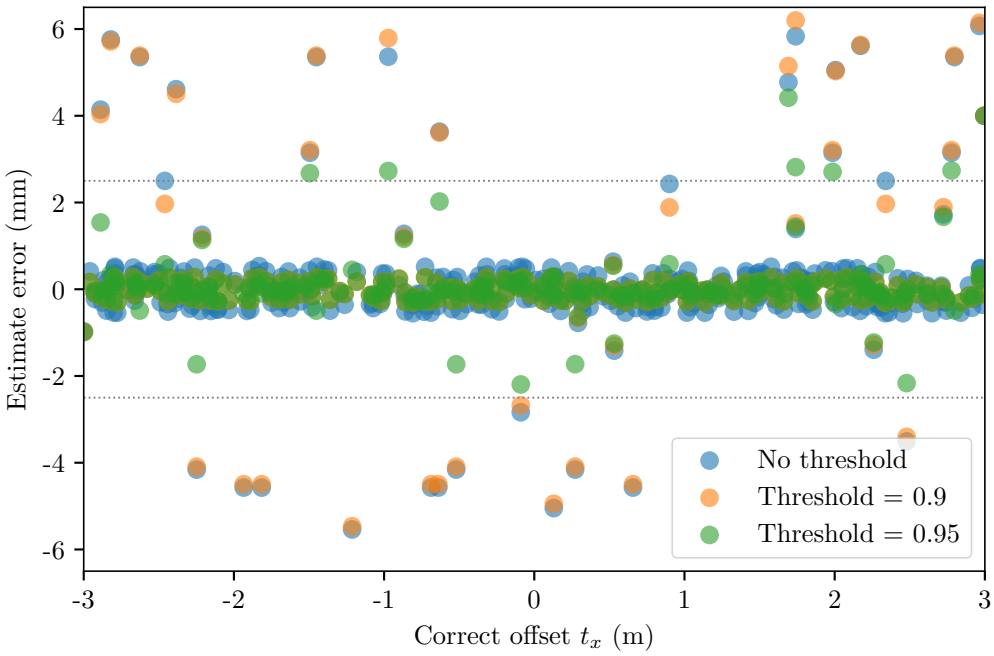


(a) Error in the position.

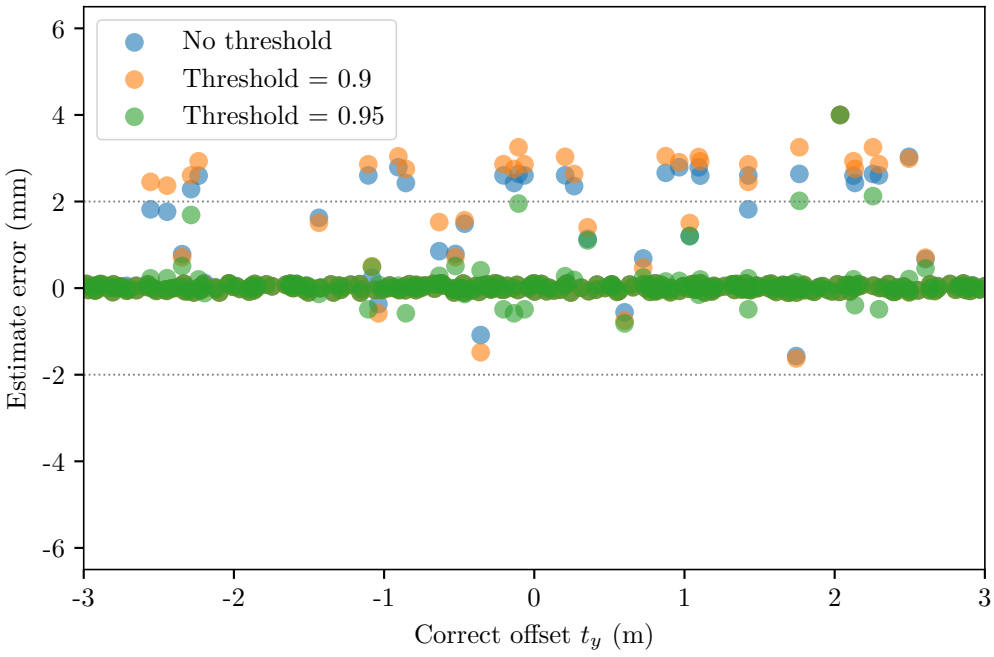


(b) Peak correlation values.

Figure 6.18: Histograms of (a) the across-track errors in the correlation peak position and (b) the values of the estimated peak for the best- and worst-case results from Figure 6.15a. The smallest mean error (less than 1 μm) occurred for $t_x = 0.42\,\mathrm{m}$, while $t_x = 2.964\,\mathrm{m}$ resulted in a mean error of 6.05 mm.

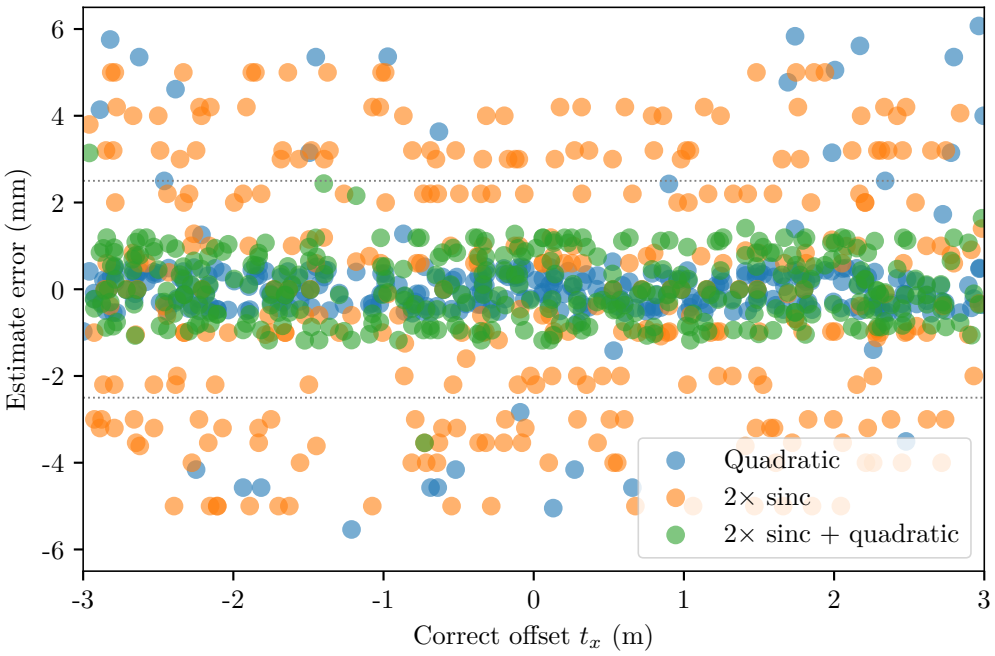


(a) Across-track error.

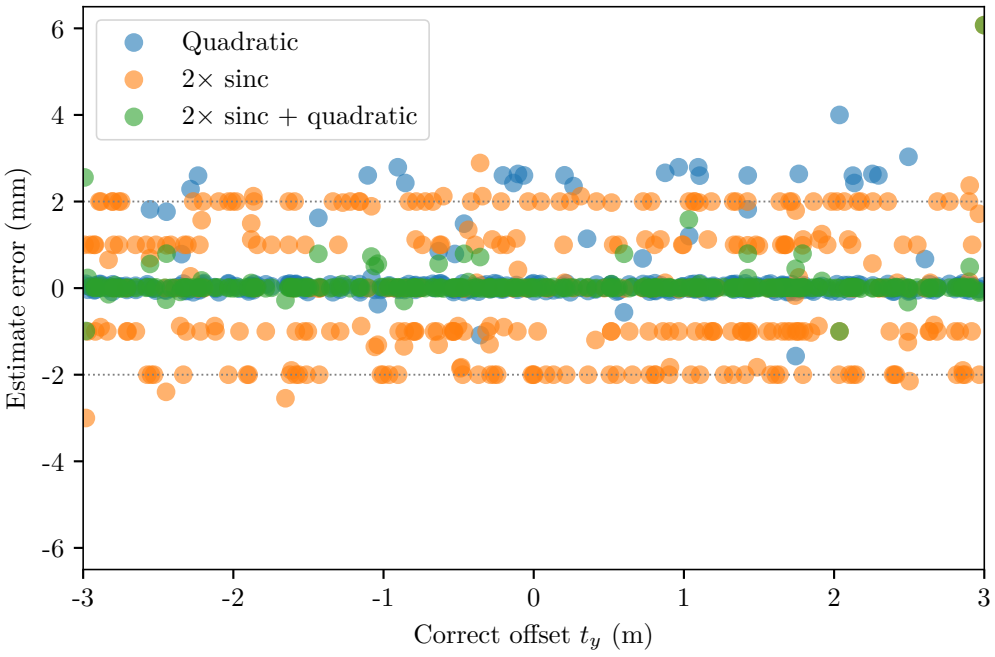


(b) Along-track error.

Figure 6.19: The effect of thresholding the positions used for the least squares estimates. Only the peak positions corresponding to a correlation value above the given threshold were used as input data for the estimation process. The target maximum misregistration of 10 % of the sonar resolution cell is marked by the dotted lines.

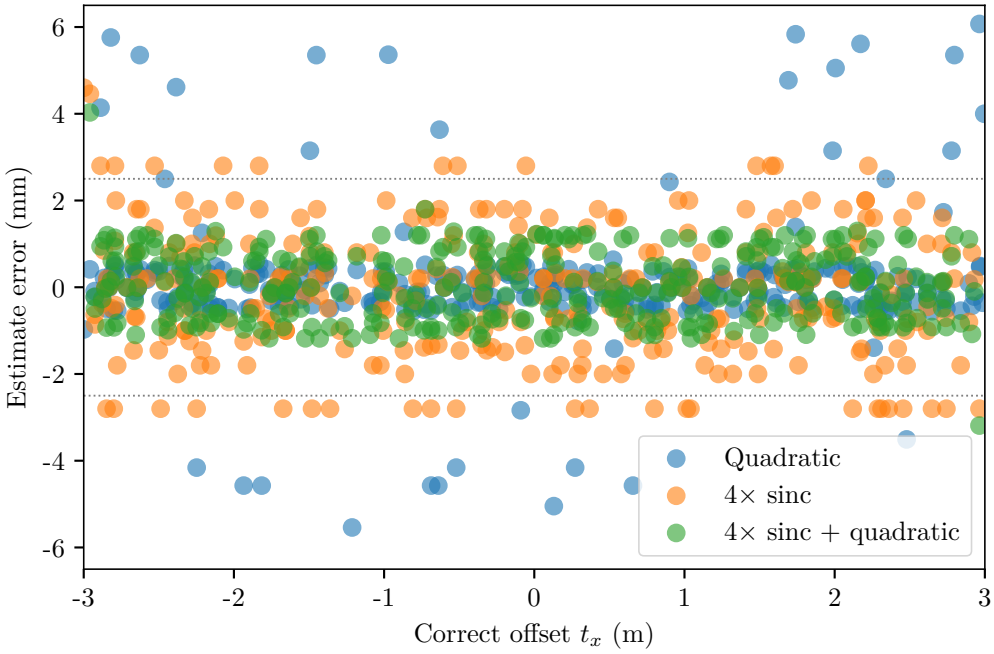


(a) Across-track error.

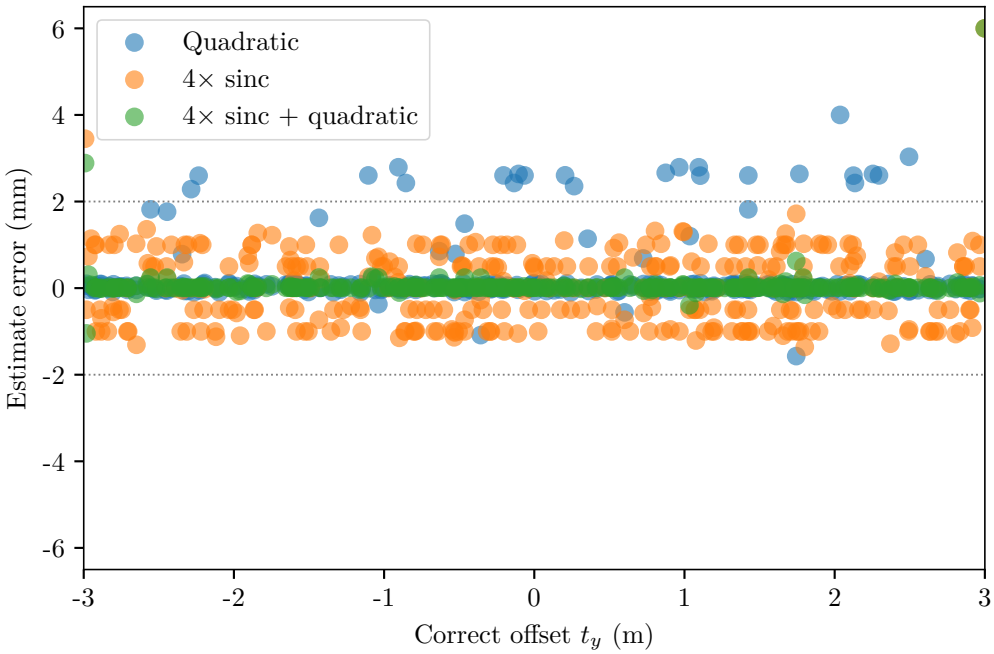


(b) Along-track error.

Figure 6.20: A comparison of the errors in the least-squares translation parameter estimates when using quadratic interpolation, 2× sinc interpolation, and 2× sinc interpolation followed by quadratic interpolation. The target maximum misregistration of 10 % of the sonar resolution cell is marked by the dotted lines.



(a) Across-track error.



(b) Along-track error.

Figure 6.21: A comparison of the errors in the least-squares translation parameter estimates when using quadratic interpolation, 4× sinc interpolation, and 4× sinc interpolation followed by quadratic interpolation. The target maximum misregistration of 10 % of the sonar resolution cell is marked by the dotted lines.

6.8.4 Affects of temporal decorrelation

The results up to this point have used imagery with no temporal decorrelation and no additive noise. To investigate the effects that these factors have on the estimation errors, the simulations were repeated for temporal coherence factors $\gamma_t = 0.9$ and $\gamma_t = 0.5$. These coherence levels can be converted into equivalent SNRs by using (3.25); they correspond to SNRs of 22 dB and 0 dB respectively. The same process as outlined in Section 6.8.1 was applied to generate a set of repeat-pass images with various translation errors in the assumed track. In this case, the errors were chosen from a uniform distribution between -2 m and 2 m in both directions. Note that the errors were not the same for each set of trials.

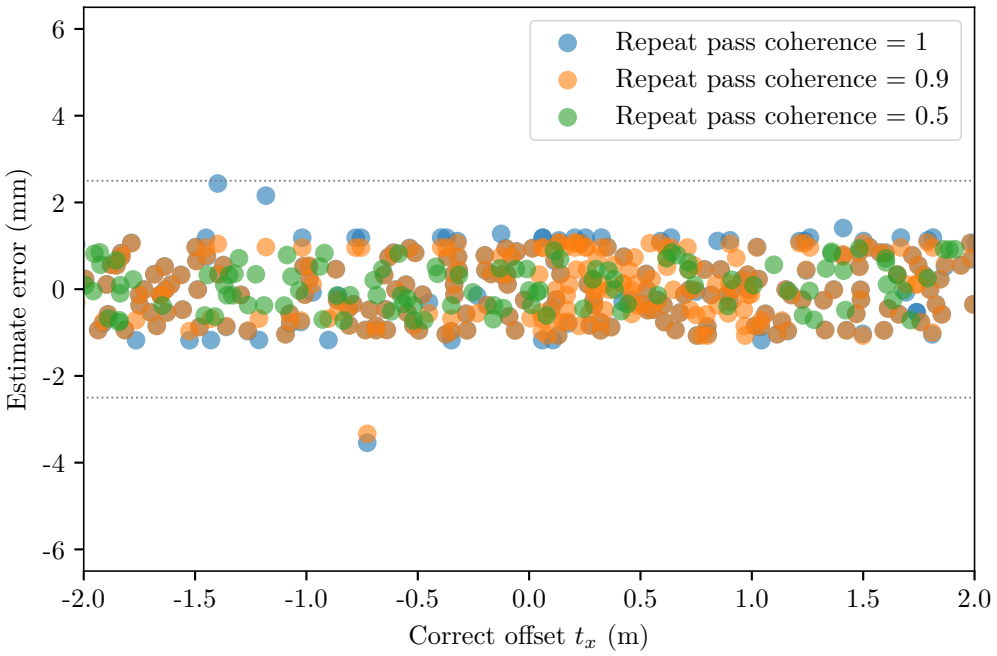
Figure 6.22 show the error in the parameter estimation (performed using the least-squares estimator) for the different levels of temporal decorrelation. The level of errors has not been increased by the loss of coherence in the repeat pass. As shown in (4.16), decorrelated speckle can be created by a linear combination of two independent noise sources. This means the temporal decorrelation can be modelled as an additive noise process with the corresponding SNR. Performing the correlation is equivalent to applying a matched filter to the repeat image to locate areas where the two images best line up. A matched filter is the optimal method of detection in the presence of additive white Gaussian noise [Crocker 1998], and allows the estimation of the parameters in cases of high temporal decorrelation (or equivalently, high levels of additive noise).

6.9 ROTATION ERRORS

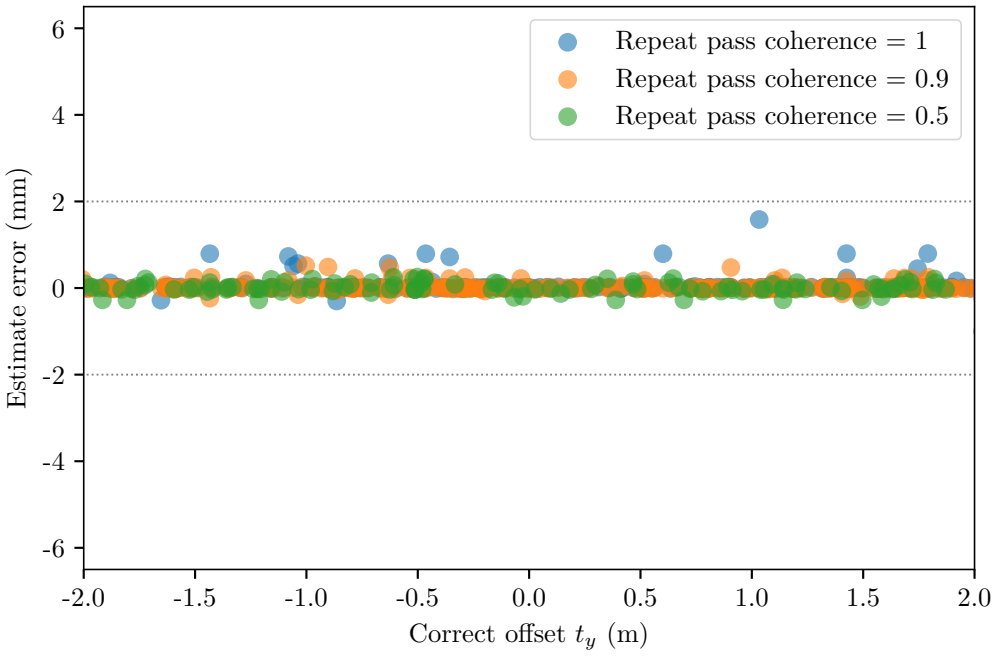
The ability of the estimators to detect rotation errors in the assumed track was examined in a similar fashion to the translation errors in the preceding section. The track followed by the sonar on the repeat pass was rotated by a range of angles between $-20°$ and $20°$. The data was then reconstructed using the backprojection algorithm with this incorrect track. The correlation was measured using 2 m by 2 m blocks in the repeat image and 8 m by 8 m blocks in the primary image, i.e., a search of ± 3 m in all directions. An example of the detected displacements is shown in Figure 6.23 for a $-3°$ heading error. Despite the displacements predicted by the model (6.18) being within the 3 m search region, no good matches were found. This implies that the rotation has led to a loss of correlation between the blocks.

6.9.1 Peak correlation magnitude versus rotation

The correlation as a function of rotation was evaluated by a simple experiment. A 142 by 142 pixel image containing pure speckle noise was created. The speckle was oversampled by the application of a Hamming window in the Fourier domain to reduce the bandwidth of the noise appropriately. A copy of the image was then taken as the



(a) Across-track error.



(b) Along-track error.

Figure 6.22: The errors in the least-squares estimation of translation parameters for different temporal decorrelation levels. A coherence factor of 0.9 corresponds to an SNR of 22 dB, and a coherence factor of 0.5 corresponds to an SNR of 0 dB. The target maximum misregistration of 10 % of the sonar resolution cell is marked by the dotted lines.

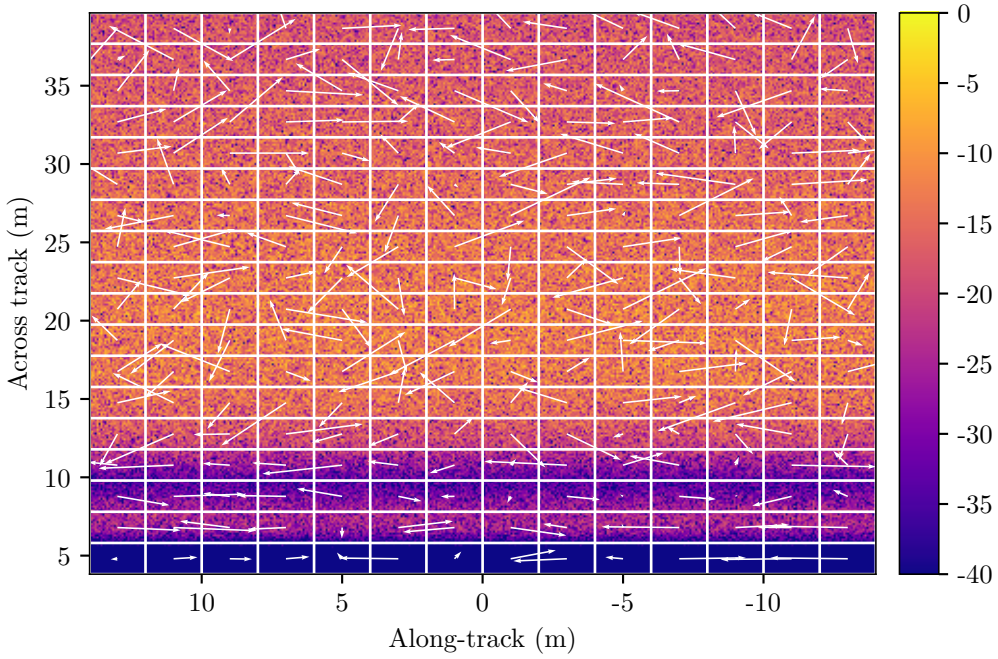


Figure 6.23: The result of direct correlation with a -3° heading error. The underlying image is the image magnitude in decibels relative to the largest value. The blocks are 2 m by 2 m and the correlation was performed with 8 m by 8 m concentric blocks of the primary pass i.e., the search was ±3 m in all directions. None of these estimated displacements fit the model.

repeat-pass image. This repeat-pass image was sinc interpolated by a factor of four in each direction prior to a bilinear interpolation to perform the rotation. A 50 by 50 pixel region was extracted from the centre of both the primary and repeat pass images in order to avoid any zero-filled areas that were introduced around the edges by the rotation process. Finally, the zero-lag correlation (equivalent to the coherence) was calculated. This was performed with 100 different speckle noise realisations at each angle.

Figure 6.24a shows the mean of the peak correlation magnitude over these 100 trials for a selection of oversampling factors for rotation angles between 0° and 90°. The shaded area shows the region that is one standard deviation either side of the mean. Figure 6.24b expands on the region between 0° and 20°. From these figures it is clear that the correlation magnitude decreases sharply as the image is rotated. In the case where no oversampling is performed, the magnitude drops below 0.5 after only 1.6° of rotation. Oversampling the image (i.e., interpolating it) slows the rate of decorrelation, albeit at the cost of a higher standard deviation in the estimates.

6.9.2 Correlation search

In order for the displacements between the image blocks to be estimated, the blocks need to be rotated to the correct orientation. This requires finding the optimal rotation between a pair of blocks. This was posed as an optimisation problem,

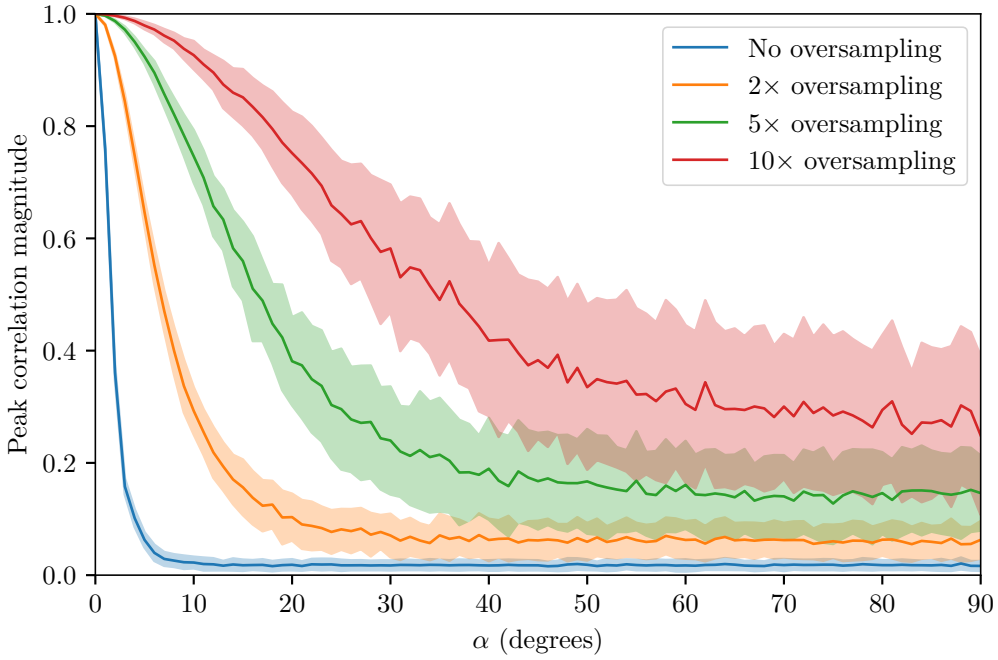
$$\hat{\beta} = \arg\min_{\beta}(1 - c(\beta)), \tag{6.32}$$

where β is the angle of rotation, and $c(\beta)$ is the peak magnitude of the correlation between the block from the primary image and the block from the repeat pass rotated by an angle β. For each of the blocks in each of the rotation trials, the L-BFGS-B minimisation algorithm [Byrd, Lu, and Nocedal 1995] was used to estimate the optimal rotation $\hat{\beta}$.

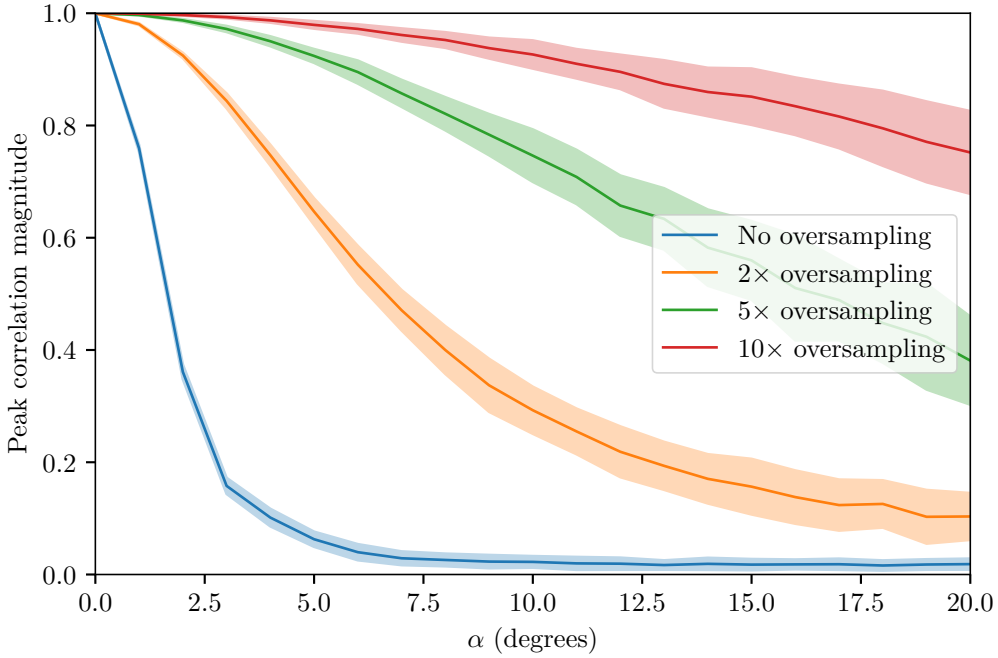
The estimated angles were then used to rotate the blocks, and the peak position of the correlation was located. The result of this procedure for the $\alpha = -3°$ heading error is shown in Figure 6.25. Rotating the blocks has resulted in the majority of the blocks yielding a displacement that follows the model.

6.9.3 Parameter estimation

The rotation parameter was estimated from the peak positions of the rotated correlations. The error in these estimates are shown in Figure 6.26a for the least squares and weighted least squares estimators, and Figure 6.26b for the optimisation and weighted optimisation estimators. Once the rotation parameter α becomes larger than four degrees, the error increases rapidly. From the model (6.18), at this point the majority of



(a) From 0° to 90° rotation.



(b) Zoomed in to show low rotation angles.

Figure 6.24: The peak correlation magnitude against image rotation. The correlation was performed over a 50×50 window. The solid lines show the mean over 100 trials, and the shaded areas show the area one standard deviation either side of the mean. The rotation of the image has caused a rapid decrease in the correlation.

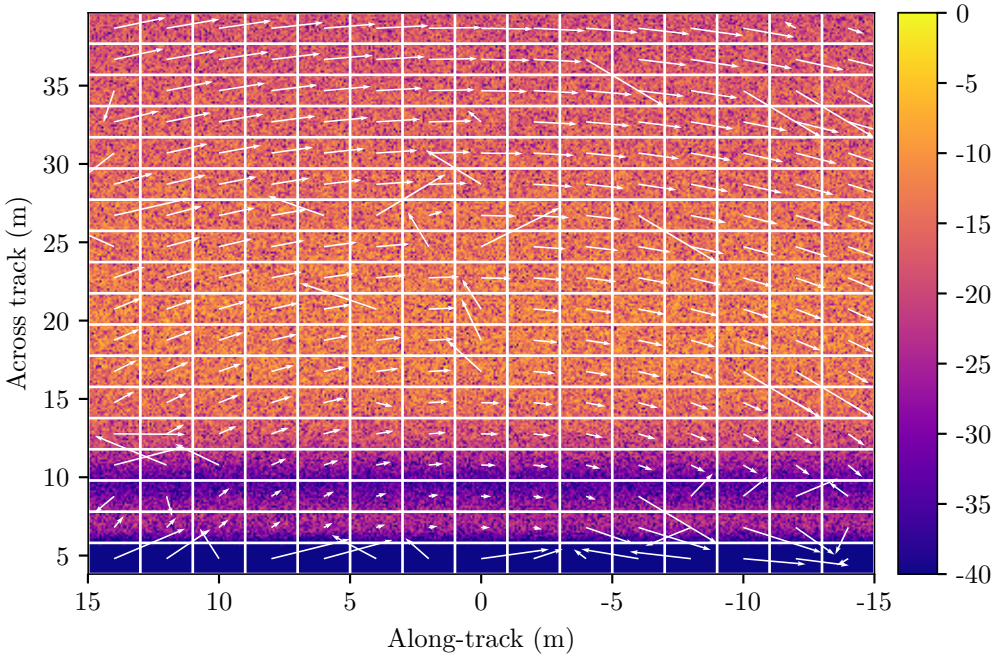


Figure 6.25: The result of a correlation search with a -3° heading error. The underlying image is the image magnitude in decibels relative to the largest value. The blocks are 2 m by 2 m and the correlation was performed with 6 m by 6 m concentric blocks of the primary pass i.e., the search was ±3 m in all directions. The majority of the observed displacements match the model although there are some blocks for which the search has failed.

the displacements become larger than 3 m and so cannot be found with the correlation. The estimators are largely similar over the region where the correlation peaks can be detected, although the weighted least squares estimator is the only one which does not give an increased error around the 3° mark.

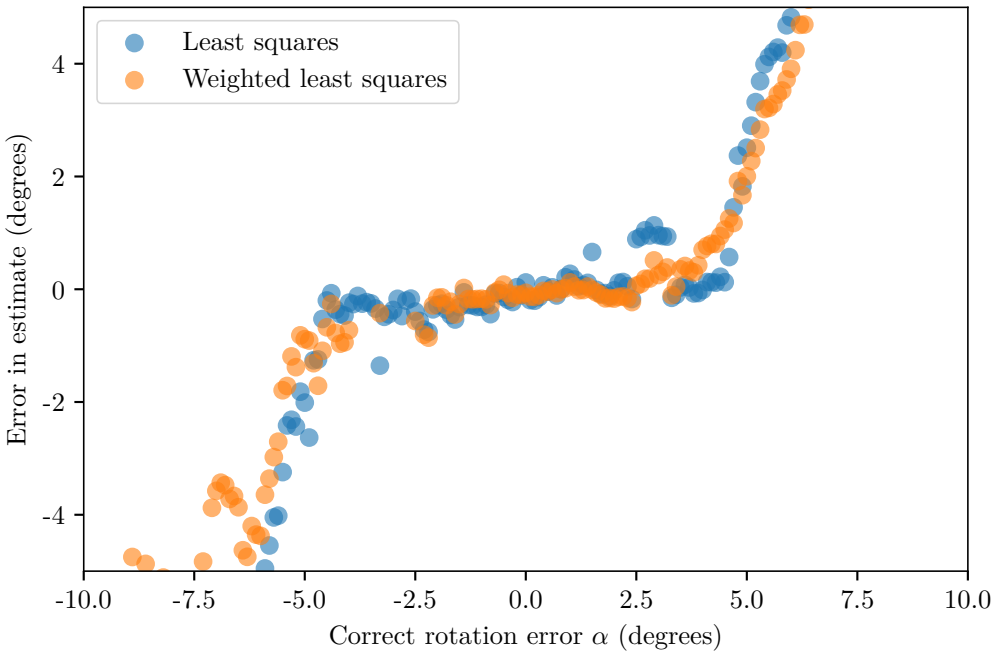
Estimating the rotation angle over a wider range of values requires an increased correlation search size. This size was increased so the peak was searched for in a region of ±5 m from the centre of the block in the repeat pass image. The same minimisation process was used to find the optimal rotation to allow a correlation estimate, the positions of the peaks were located, and the four estimators were used to find the rotation parameter. The least squares results are shown in Figure 6.27a and the optimisation results are in Figure 6.27b. The parameter is now able to be estimated in the region from −8° to 8°. As with the previous example, the four estimators are similar in terms of the level of error.

6.10 DISCUSSION

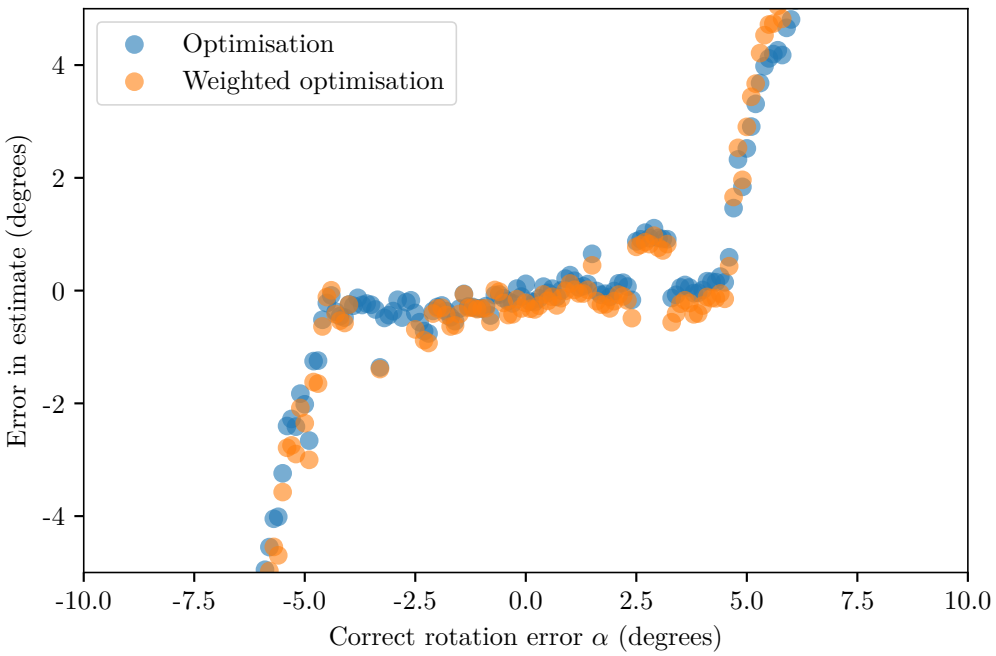
Registering repeat-pass images first requires them to be reconstructed onto a common grid. A simple method of achieving this was presented. A study of a SAS system with high-quality navigation data showed that the misregistration errors were too large to achieve a suitable coherence level for change detection or repeat-pass interferometry. This indicates the need for data-driven registration to correct for errors in the assumed path of the sonar.

A simple model of the heading and translation errors in a nominally straight track was presented. From this, the expected displacement from its true location of a point in the scene due to these can be calculated. A block-based correlation technique was suggested as a method of measuring the displacement between two images. Two estimation techniques were then suggested for taking the measured displacement field and estimating the corresponding model parameters.

A number of simulated datasets with errors in the assumed track were used to evaluate the performance of this registration algorithm. One of the key findings was the need for accurate estimation of the location of the peak in the correlation of two image blocks. Although quadratic interpolation is a simple and efficient method of estimating this, it has a bias which in some cases led to a significant error in the final estimated position. Sinc oversampling the images prior to using the quadratic interpolation was shown to reduce this maximum error to below the desired threshold for minimising misregistration error. Boucher and Hassab [1981] suggest applying a window to reduce the bias in the quadratic interpolator. However, this results in a trade-off: windowing the autocorrelation will broaden its peak which may worsen the accuracy of the estimate.

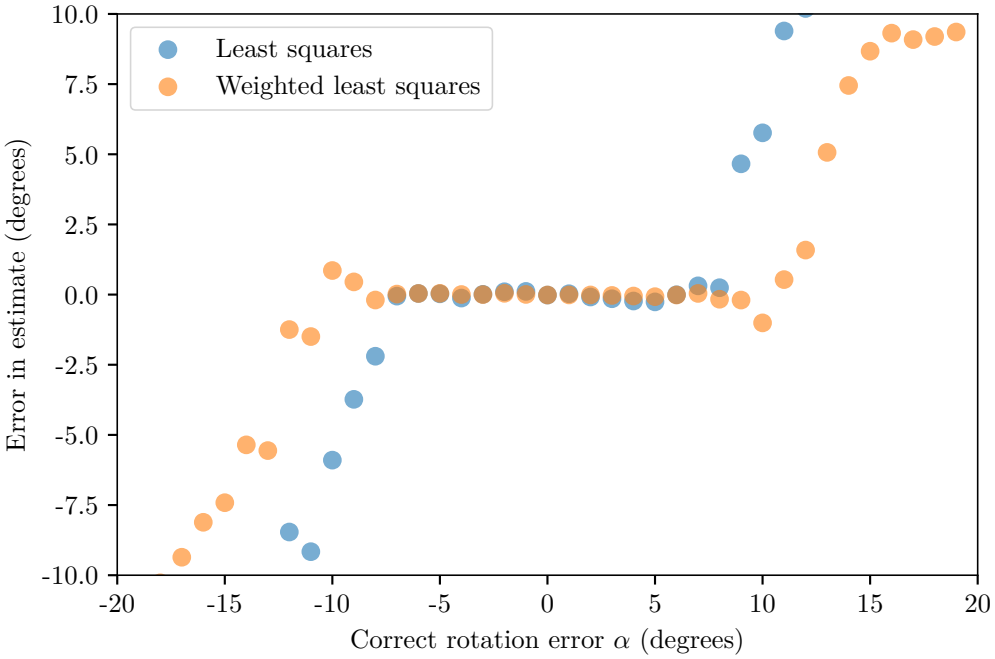


(a) Least squares and weighted least squares.

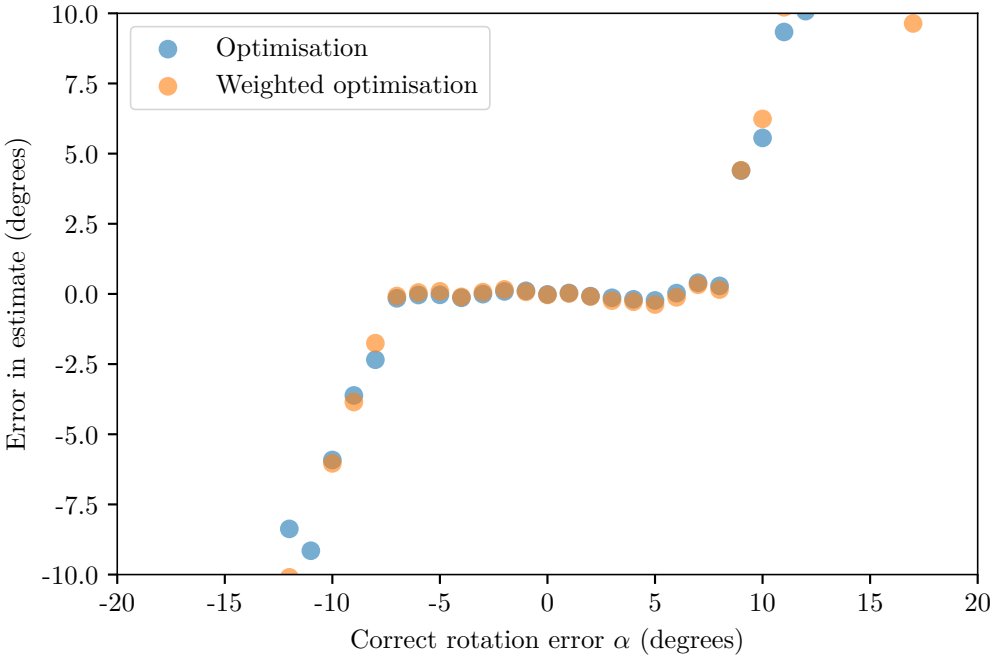


(b) Optimisation and weighted optimisation.

Figure 6.26: The error in the estimated rotation parameter α when the correlation is performed with 2 m by 2 m blocks and a search of 3 m in all directions. As the angle gets larger, the displacements predicted by the model (6.18) are too large to be found by this search and so the parameter estimation fails.



(a) Least squares and weighted least squares.



(b) Optimisation and weighted optimisation.

Figure 6.27: The error in the estimated rotation parameter α when the correlation is performed with 2 m by 2 m blocks and a search of 5 m in all directions. Compared to Figure 6.26, the larger search region allows the estimation of larger rotation parameters.

The presented technique worked accurately when detecting translation errors in the assumed track. Using correlation as the method of calculating the displacement between images is equivalent to applying a matched filter. As such, this is the optimal approach for detection of a signal in additive white Gaussian noise. As a temporal decorrelation can be modelled as an additive process, this results in the correlation method being robust in both low-coherence and low-SNR scenarios.

Correlation does not work directly with rotated images. With no oversampling, it was shown that the correlation peak becomes negligible with less than 2° degrees of rotation. When looking for heading errors, finding the displacement of a repeat-pass block first requires it to be rotated into alignment with the primary pass. As the rotation angle this requires is not known, a search method is required. A minimisation problem was formulated and shown to be effective at finding this angle. With the repeat-pass blocks rotated, the displacement field was able to be calculated and the heading error estimated.

The accuracy of the correlation method is dictated by the autocorrelation of the impulse response of the system. This is determined by the signal bandwidth in the across-track direction, and by the width and shape of the transmitting and receiving apertures in the along-track direction. Careful design of the system allows the width of the autocorrelation peak to be minimised, thus minimising errors in the calculated repeat-pass displacement field.

The search size of the block correlation technique is a critical parameter in the registration process. For translation errors, the search size needs to be large enough to encompass the largest expected error. The displacement due to a heading error varies with the range of the block from the sonar. The search size in this case needs to be predicted from the presented model, taking into account the largest expected heading error and the maximum range of the sonar. As the overall displacement is a linear combination of the translation and heading components, the final search size required to estimate the displacement can be calculated from the addition of the maximum translation error and the largest expected displacement due to a heading error.

Chapter 7

CONCLUSIONS

Simulation of repeat-pass data is useful as it provides a known ground-truth with which to compare the results of any registration algorithm. The rough facet scattering model uses complex Gaussian noise to generate realisations of statistically rough facets. It was shown that care is needed to make these realisations identical for repeated passes of the scene. With suitable modifications, the existing simulator was able to generate repeat-pass images.

For realistic data, areas of the simulated scene need to exhibit temporal decorrelation between the passes. A process based on a linear combination of noise was developed for the generation of partially correlated complex Gaussian noise. Using this correlated noise as the basis of generating the rough facet realisations results in the repeat-pass coherence of the facets dropping by the specified amount. This process was implemented in the simulator, and a number of scenarios were simulated to demonstrate its operation. These results were validated against the published coherence models for speckle images by use of the Kolmogorov-Smirnov goodness-of-fit test.

Comparison of repeat-pass images requires them to be aligned on a common grid. The choice of coordinate system is arbitrary provided the same coordinates are used for all images. However, it makes sense for it to have some relationship to the paths followed by the sonar. A method of generating a suitable coordinate system based on an idealised linear path was developed.

In an ideal scenario, the navigation data will be sufficiently accurate to minimise registration errors between the images. A study was undertaken of some data acquired from a ship-mounted sonar system which was instrumented with a real-time kinematic GPS. Despite the high-quality navigation data available, the alignment of these images still exhibited significant misregistration errors. An estimate of the misregistration for a portion of the image was obtained via correlation despite the low temporal coherence of the images. This shows that correlation is a viable technique for detecting the misregistration.

Correcting the assumed tracks requires a model of how the track errors affect the registration of the images. A model taking into account both translation and heading

errors in a track was presented, and the relationship between these parameters and the displacement of a given area of the image was developed. Calculating this displacement allows an estimate of the track error parameters to be generated.

A block-based correlation method was presented to estimate the displacement of each portion of a repeat pass image with respect to the primary pass image. This was tested with simulated SAS images and shown to be robust for images corrupted by a low SNR or low temporal coherence. Correlation was shown to rapidly degrade with the rotation of a block. In order to estimate the displacement of an image resulting from a track with a heading error, each repeat-pass block needs to be rotated to align it with the corresponding area of the primary image. Finding the optimal rotation angle to achieve this was posed as a minimisation problem. Following the estimation of this optimal angle, correlating the rotated blocks allowed the calculation of the displacement between the images, and from this the parameters of the track error were able to be estimated.

A number of guidelines for maximising the ability to perform this data-driven registration can be drawn from this research. The accuracy of the block correlation technique in detecting the displacement between images is dependent on the width of the peak in the autocorrelation of the images. This peak width is dependent on the transmitted signal bandwidth, and on the size and shape of the apertures used for transmitting and receiving. Designing the system to minimise the peak width allows a higher accuracy.

The interpolation method used to estimate the position of the peak correlation has a significant impact on the accuracy of the results. Oversampling the correlation image prior to applying the interpolator reduces the errors to a sufficiently low level to allow accurate registration. The interpolation method chosen should be evaluated for a given system.

The appropriate correlation search size depends on the expected maximum errors in the assumed track. The track error model allows the maximum displacement to be calculated for these maximum errors. From this, a suitable search size can be determined.

The lacunarity and coherence calculations presented in this thesis require repeated evaluation over a sliding window. This is computationally expensive to implement directly. The use of integral images, a concept developed in the computer graphics field, has previously been shown to improve the efficiency of lacunarity calculations. This has been extended to the calculation of coherence.

7.1 SUGGESTIONS FOR FUTURE RESEARCH

Throughout this research a number of potential improvements and extensions to the presented work have been identified:

Long-tailed speckle statistics: The circular Gaussian multiplicative noise model of speckle ('fully developed speckle') is sufficient for featureless seafloors. In other cases, this model is not an accurate description of the speckle. Instead of being Rayleigh-distributed, the magnitude of the imagery takes on a longer-tailed distribution. A common alternative model for the magnitude is the K-distribution [Lyons and Abraham 1999]. For example, the statistics of a rippled seafloor have been shown to follow a K-distribution based on the slope of the ripples [Lyons, Abraham, and Johnson 2010]. Modifying the simulator to generate images with K-statistics would allow evaluation of the registration procedures for a wider variety of scenarios. Some research into how to simulate decorrelated sand ripples is required.

Piecewise-linear track model: The track error model presented assumes a linear track. This could be extended into a piecewise-linear model to represent a non-linear track. Research is needed into which portions of the track would affect the displacement of a given point in the scene.

Validation with real-world data: The presented registration algorithm works with simulated data but has not yet been tested with real-world data. The availability of suitable data is the key here, especially in regards to the ground truth required for validation of the results. One option would be to use data from a single pass and perturb the recorded track to generate a repeat-pass image. The performance of the algorithm can then be compared to these known errors. An alternative would be to use a bland seafloor area bordered by features such as rocks. After running the registration algorithm on the blank areas, the alignment of the adjacent features can be used to evaluate the accuracy of the registration. A similar approach could be used to test the algorithm in the presence of temporal decorrelation.

Improved rotation estimation: The rotation of each block to allow the correlation to detect the displacement used a minimisation approach. A more computationally efficient approach should be investigated.

Error estimation during reconstruction: The backprojection algorithm can reconstruct a small portion of an image (a region of interest). The reconstruction of the repeat-pass image could be divided into a number of these regions and sequentially reconstructed. Estimating the track errors from each region as it is reconstructed might allow the subsequent blocks to be more accurately processed. This may be more effi-

cient than registering the image after reconstruction.

Combination with computer vision techniques: Wang and Hayes [2017] have approached the registration of bland seafloors using computer vision feature-matching techniques such as SIFT. They have demonstrated on an ideal simulated scene that this can generate sufficient correspondences to perform an accurate registration. Combining this approach with the correlation techniques presented here may improve the efficiency of the registration process.

Use of features: A largely bland seafloor may have occasional features such as rocks, debris, etc. These can provide stable reference points as they have a distinct shape and are significantly less susceptible to temporal decorrelation compared to the seafloor. The correlation of such features is likely to be more accurate than that of the surrounding seafloor, and as such should carry a higher weighting in the algorithm. A method of identifying such features (e.g., based on the lacunarity of the scene) is required to allow this weighting to be determined.

Applicability to incoherent change detection: This thesis has concentrated on the use of this registration algorithm with coherent change detection. Although the required registration accuracy is less stringent for incoherent change detection, it is still desirable to be as accurate as possible. Some preliminary tests have shown that the presented algorithm works with incoherent (magnitude-only) data, although the possible loss of accuracy arising from the lack of phase information has not yet been explored.

Affects of autofocus: Small-scale wobbles in the sonar path result in blurring in the reconstructed imagery. A number of autofocus algorithms exist to remove these effects. This modification may have an affect on the estimation of the track error parameters, especially if the blurring in one image is different to the other. The interaction between the autofocus and registration algorithms should be studied; it may also be possible to guide one based on the findings of the other.

Sensor fusion: The estimated track found by the registration algorithm could be augmented with readings from inertial measurement units, Doppler velocity logs, or other sources of data via a sensor fusion process such as extended Kalman filtering.

Appendix A

REPRESENTING ROTATIONS

A rotation is characterised by an axis of rotation,

$$\mathbf{a} = [a_x, a_y, a_z]^{\mathrm{T}}, \tag{A.1}$$

and an angle of rotation, θ. Since only the direction of $\mathbf{a}$ is important, it can be normalised to give

$$\hat{\mathbf{a}} = [\hat{a}_x, \hat{a}_y, \hat{a}_z]^{\mathrm{T}}. \tag{A.2}$$

A positive rotation is defined as one that matches the right-hand rule, that is, one which appears counter-clockwise when looking along the axis towards the origin. Such a rotation is illustrated in Figure A.1. It is possible to choose the rotations to match the left-hand rule, i.e., a positive rotation being clockwise when looking towards the origin. This essentially just flips the sign of the angle of rotation. However, many mathematical constructs (e.g., the cross-product) are based on the right-hand rule, and defining the rotation opposing this can lead to subtle issues.

For convenience, this thesis uses rotation matrices to represent a rotation. This appendix details the derivation and application of these matrices.

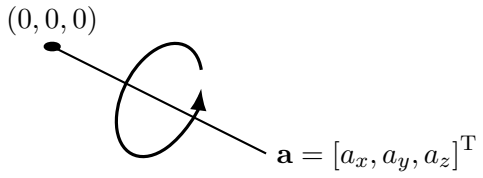


Figure A.1: A positive rotation about an axis $\mathbf{a}$.

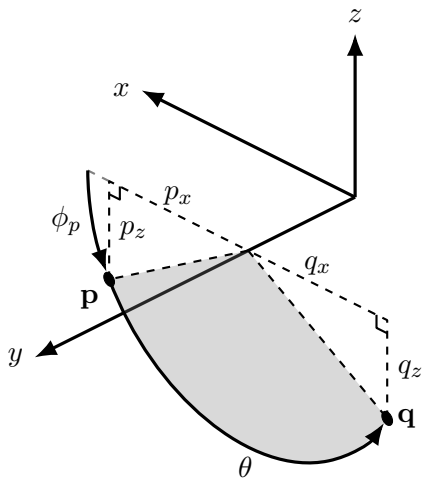


Figure A.2: Rotating a point **p** around the y-axis to **q**.

A.1 ROTATION MATRICES

A rotation can be specified by a *rotation matrix*, a three-by-three matrix

$$\mathbf{R} = \begin{bmatrix} R_{xx} & R_{xy} & R_{xz} \\ R_{yx} & R_{yy} & R_{yz} \\ R_{zx} & R_{zy} & R_{zz} \end{bmatrix}. \tag{A.3}$$

A point $\mathbf{p} = [p_x, p_y, p_z]^{\mathrm{T}}$ can then be rotated via matrix multiplication to a new point $\mathbf{q} = \mathbf{R}\mathbf{p}$ with the coordinates

$$\mathbf{q} = \begin{bmatrix} R_{xx}p_x + R_{xy}p_y + R_{xz}p_z \\ R_{yx}p_x + R_{yy}p_y + R_{yz}p_z \\ R_{zx}p_x + R_{zy}p_y + R_{zz}p_z \end{bmatrix}. \tag{A.4}$$

A.2 ROTATION ABOUT THE AXES

The basic rotations are those around the axes of the coordinate system. By combining these, any arbitrary rotation can be applied. The following three sections derive the rotation matrices for these basic rotations.

A.2.1 Y axis

Figure A.2 shows the geometry of point **p** rotated around the y-axis (i.e., $\hat{\mathbf{a}} = [0, 1, 0]^{\mathrm{T}}$) by an angle θ to create a new point **q**. The initial point **p** is a distance $\rho_y = \sqrt{p_x^2 + p_z^2}$ from the y-axis at an angle ϕ_p (measured from the positive x-axis). We can define its x- and z-coordinates as

$$p_x = \rho_y \cos\phi_p, \tag{A.5}$$

$$p_z = -\rho_y \sin\phi_p. \tag{A.6}$$

Note that the negative sign in the z-coordinate arises from the definition of the rotation: it is negative in the first two quadrants of ϕ_p and positive in the last two. Similarly, the x-coordinate of the rotated point can be defined as

$$\begin{aligned} q_x &= \rho_y \cos(\phi_p + \theta), \\ &= \rho_y \cos\phi_p \cos\theta - \rho_y \sin\phi_p \sin\theta, \\ &= p_x \cos\theta + p_z \sin\theta. \end{aligned} \tag{A.7}$$

Using the same procedure for the z-coordinate of the rotated point gives

$$\begin{aligned} q_z &= -\rho_y \sin(\phi_p + \theta), \\ &= -\rho_y \cos\phi_p \sin\theta - \rho_y \sin\phi_p \cos\theta, \\ &= -p_x \sin\theta + p_z \cos\theta. \end{aligned} \tag{A.8}$$

Applying these equations to (A.4) shows that the rotation matrix for a rotation by θ around the y-axis is

$$\mathbf{R_y} = \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix}. \tag{A.9}$$

A.2.2 X axis

Figure A.3 shows the geometry of rotating a point around the x-axis. The angle of the initial point is measured from the positive y-axis, and its y- and z-coordinates are given by

$$p_y = \rho_x \cos\phi_p, \tag{A.10}$$

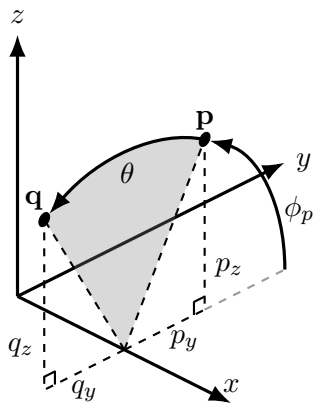


Figure A.3: Rotating a point **p** around the x-axis to **q**.

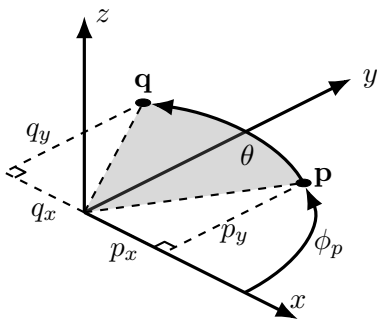


Figure A.4: Rotating a point **p** around the z-axis to **q**.

and

$$p_z = \rho_x \sin\phi_p, \tag{A.11}$$

where $\rho_x = \sqrt{p_y^2 + p_z^2}$ is the distance between **p** and the x-axis. The coordinates of the rotated point are then

$$\begin{aligned} q_y &= \rho_x \cos(\phi_p + \theta), \\ &= \rho_x \cos\phi_p \cos\theta - \rho_x \sin\phi_p \sin\theta, \\ &= p_y \cos\theta - p_z \sin\theta, \end{aligned} \tag{A.12}$$

and

$$\begin{aligned} q_z &= \rho_x \sin(\phi_p + \theta), \\ &= \rho_x \cos\phi_p \sin\theta + \rho_x \sin\phi_p \cos\theta, \\ &= p_y \sin\theta + p_z \cos\theta. \end{aligned} \tag{A.13}$$

This leads to the rotation matrix

$$\mathbf{R_x} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}. \tag{A.14}$$

A.2.3 Z axis

The geometry of a rotation about the z-axis is shown in Figure A.4. With the initial angle being measured from the positive x-axis, and the distance between the z-axis and the point $\rho_z = \sqrt{p_x^2 + p_y^2}$, the x- and y-coordinates of **p** are

$$p_x = \rho_z \cos\phi_p, \tag{A.15}$$

and

$$p_y = \rho_z \sin\phi_p. \tag{A.16}$$

The coordinates of the rotated point $\mathbf{q}$ are

$$\begin{aligned} q_x &= \rho_z \cos(\phi_p + \theta), \\ &= \rho_z \cos\phi_p \cos\theta - \rho_z \sin\phi_p \sin\theta, \\ &= p_x \cos\theta - p_y \sin\theta, \end{aligned} \tag{A.17}$$

and

$$\begin{aligned} q_y &= \rho_z \sin(\phi_p + \theta), \\ &= \rho_z \cos\phi_p \sin\theta + \rho_z \sin\phi_p \cos\theta, \\ &= p_x \sin\theta + p_y \cos\theta. \end{aligned} \tag{A.18}$$

Therefore, a rotation about the z-axis has the rotation matrix

$$\mathbf{R_z} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{A.19}$$

A.3 COMBINING ROTATIONS

The order of rotations matters: a rotation around the y-axis followed by a rotation around the z-axis does not give the same result as a rotation around the z-axis followed by a rotation around the y-axis. Suppose there are N rotations, represented by the rotation matrices $\mathbf{R_1}, \mathbf{R_2}, \ldots, \mathbf{R_{N-1}}, \mathbf{R_N}$ which are to be applied in that order to a point $\mathbf{p}$. The result of the first rotation is

$$\mathbf{q} = \mathbf{R_1 p}. \tag{A.20}$$

The result of the second rotation is then

$$\begin{aligned} \mathbf{r} &= \mathbf{R_2 q}, \\ &= \mathbf{R_2 R_1 p}. \end{aligned} \tag{A.21}$$

More generally, the result of all N rotations is

$$\mathbf{s} = \mathbf{R_N R_{N-1}} \ldots \mathbf{R_2 R_1 p}. \tag{A.22}$$

Therefore, the individual rotation matrices can be combined into an overall matrix,

$$\mathbf{R} = \prod_{n=0}^{N-1} \mathbf{R_{N-n}}, \tag{A.23}$$

which can be applied to a point in a single operation

$$\mathbf{s} = \mathbf{R p}. \tag{A.24}$$

A.4 ROTATION ABOUT AN ARBITRARY AXIS

A rotation θ about an arbitrary axis can be performed as a series of five steps:

1. Rotate around the z-axis to put the axis of rotation in the x-z plane.
2. Rotate around the y-axis to align the axis of rotation and the z-axis.
3. Perform the rotation by θ around the axis of rotation/z-axis.
4. Undo the y-axis rotation to move the axis of rotation back into the x-z plane.
5. Undo the z-axis rotation to move the axis of rotation back to its original position.

The geometry of the rotation into the x-z plane is shown in Figure A.5a with α being the neccessary angle of rotation. From this,

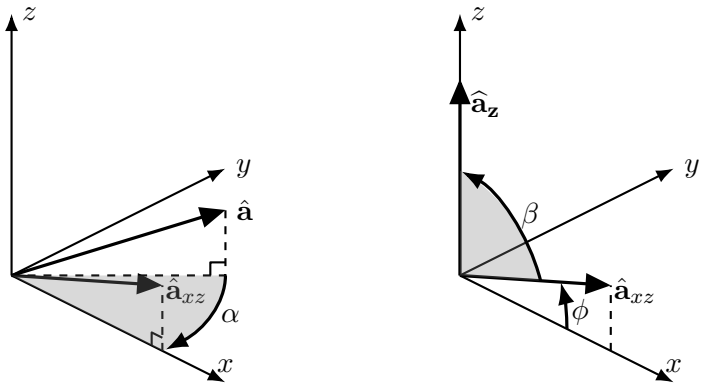
$$\cos\alpha = \frac{\hat{a}_x}{\sqrt{\hat{a}_x^2 + \hat{a}_y^2}}, \tag{A.25}$$

and

$$\sin\alpha = \frac{\hat{a}_y}{\sqrt{\hat{a}_x^2 + \hat{a}_y^2}}. \tag{A.26}$$

Since this is a negative rotation around the z-axis, the corresponding rotation matrix is

$$\mathbf{R_1} = \begin{bmatrix} \dfrac{\hat{a}_x}{\sqrt{\hat{a}_x^2 + \hat{a}_y^2}} & \dfrac{\hat{a}_y}{\sqrt{\hat{a}_x^2 + \hat{a}_y^2}} & 0 \\ -\dfrac{\hat{a}_y}{\sqrt{\hat{a}_x^2 + \hat{a}_y^2}} & \dfrac{\hat{a}_x}{\sqrt{\hat{a}_x^2 + \hat{a}_y^2}} & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{A.27}$$



(a) Rotate into the $x-z$ plane. (b) Rotate to align with the z-axis.

Figure A.5: Steps to align an arbitrary axis of rotation $\hat{\mathbf{a}}$ with the z-axis.

The geometry of the rotation to align $\hat{\mathbf{a}}_{xz}$ with the z-axis is shown in Figure A.5b. This gives

$$\begin{aligned}\cos\beta &= \cos(\pi/2-\phi),\\ &= \sin(\phi),\\ &= \frac{\hat{a}_z}{\sqrt{\hat{a}_x^2+\hat{a}_y^2+\hat{a}_z^2}},\end{aligned} \tag{A.28}$$

and

$$\begin{aligned}\sin\beta &= \sin(\pi/2-\phi),\\ &= \cos(\phi),\\ &= \frac{\sqrt{\hat{a}_x^2+\hat{a}_y^2}}{\sqrt{\hat{a}_x^2+\hat{a}_y^2+\hat{a}_z^2}}.\end{aligned} \tag{A.29}$$

As this is a negative rotation about the y-axis, the corresponding rotation matrix is

$$\mathbf{R_2} = \begin{bmatrix} \dfrac{\hat{a}_z}{\sqrt{\hat{a}_x^2+\hat{a}_y^2+\hat{a}_z^2}} & 0 & -\dfrac{\sqrt{\hat{a}_x^2+\hat{a}_y^2}}{\sqrt{\hat{a}_x^2+\hat{a}_y^2+\hat{a}_z^2}} \\ 0 & 1 & 0 \\ \dfrac{\sqrt{\hat{a}_x^2+\hat{a}_y^2}}{\sqrt{\hat{a}_x^2+\hat{a}_y^2+\hat{a}_z^2}} & 0 & \dfrac{\hat{a}_z}{\sqrt{\hat{a}_x^2+\hat{a}_y^2+\hat{a}_z^2}} \end{bmatrix}. \tag{A.30}$$

The desired rotation about $\hat{\mathbf{a}}$ is now equivalent to a positive rotation around the z-axis, meaning its rotation matrix is

$$\mathbf{R}_\theta = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{A.31}$$

The overall rotation matrix is given by

$$\mathbf{R} = \mathbf{R_1^T R_2^T R_\theta R_2 R_1}. \tag{A.32}$$

Performing these multiplications[1] and simplifying using $\sqrt{\hat{a}_x^2+\hat{a}_y^2+\hat{a}_z^2}=1$ gives the rotation matrix for rotation about an arbitrary axis as

$$\mathbf{R} = \begin{bmatrix} \hat{a}_x^2 b+\cos\theta & \hat{a}_x\hat{a}_y b-\hat{a}_z\sin\theta & \hat{a}_x\hat{a}_z b+\hat{a}_y\sin\theta \\ \hat{a}_x\hat{a}_y b+\hat{a}_z\sin\theta & \hat{a}_y^2 b+\cos\theta & \hat{a}_y\hat{a}_z b-\hat{a}_x\sin\theta \\ \hat{a}_x\hat{a}_z b-\hat{a}_y\sin\theta & \hat{a}_y\hat{a}_z b+\hat{a}_x\sin\theta & \hat{a}_z^2 b+\cos\theta \end{bmatrix}, \tag{A.33}$$

where $b = 1-\cos\theta$.

[1] A computer algebra system like SymPy or Sage is useful for this part.

A.5 INVERTING A ROTATION

The entries of the three basic rotation matrices are all 0, 1, $\cos\theta$ or $\sin\theta$. If it is desired to generate an inverse of the rotation matrix, i.e., rotate by $-\theta$ instead of θ, then the cosine term remains unchanged as $\cos\theta$ is an even function. Similarly, the sine term is replaced by $-\sin\theta$ since it is odd function. By inspection, making this replacement in all three basic rotation matrices results in the transpose of the original matrix. A similar replacement in the general rotation matrix (A.33) results in the same outcome, i.e., a rotation applied by a rotation matrix $\mathbf{R}$ can be inverted by the rotation matrix $\mathbf{R}^{\mathrm{T}}$.

Appendix B

THE FRAUNHOFER APPROXIMATION

In a homogeneous medium, radiated energy spreads spherically from a idealised point source. Consider a source located at $\mathbf{x}_s$. The value of the field ψ at some position $\mathbf{x}$ is a function of the distance between the source and the measurement point, i.e.,

$$\psi(\mathbf{x}) = f(r), \tag{B.1}$$

where $r = |\mathbf{x} - \mathbf{x}_s|$ is the range. This can be related to a measurement of the field at a different point $\mathbf{x}_0$. If the offset vector $\mathbf{x}' = \mathbf{x} - \mathbf{x}_0$ is the vector between the points, then the range can be rewritten as

$$r = \left|(\mathbf{x}_0 - \mathbf{x}_s) + \mathbf{x}'\right|. \tag{B.2}$$

Using the fact that[1] $|\mathbf{a} + \mathbf{b}| = \sqrt{|\mathbf{a}|^2 + |\mathbf{b}|^2 + 2(\mathbf{a} \cdot \mathbf{b})}$, the range can be calculated as

$$r = \sqrt{|\mathbf{x}_0 - \mathbf{x}_s|^2 + |\mathbf{x}'|^2 + 2(\mathbf{x}_0 - \mathbf{x}_s) \cdot \mathbf{x}'}. \tag{B.3}$$

This can be rewritten as

$$\begin{aligned} r &= \sqrt{r_0^2 \left(1 + \frac{|\mathbf{x}'|^2 + 2(\mathbf{x}_0 - \mathbf{x}_s) \cdot \mathbf{x}'}{r_0^2}\right)}, \\ &= r_0 \sqrt{1 + \frac{|\mathbf{x}'|^2 + 2(\mathbf{x}_0 - \mathbf{x}_s) \cdot \mathbf{x}'}{r_0^2}}, \end{aligned} \tag{B.4}$$

where $r_0 = |\mathbf{x}_0 - \mathbf{x}_s|$ is the range between the source and the second measurement point. Defining $\rho^2 = |\mathbf{x}'|^2 + 2(\mathbf{x}_0 - \mathbf{x}_s) \cdot \mathbf{x}'$ allows this to be simplified to

$$r = r_0 \sqrt{1 + \frac{\rho^2}{r_0^2}}. \tag{B.5}$$

[1] This is the law of cosines using vector notation and the definition of the dot product.

Figure B.1: Shown here for a simple 2-dimensional case, the Fraunhofer approximation projects the offset vector $\mathbf{x}'$ onto the vector between the source $\mathbf{x}_s$ and the second point $\mathbf{x}_0$ to estimate the range to $\mathbf{x}$.

Applying the binomial expansion $\sqrt{1+u} = 1 + u/2 - u^2/8 + \cdots$ gives the range as

$$\begin{aligned} r &= r_0\left[1 + \frac{\rho^2}{2r_0^2} - \frac{\rho^4}{8r_0^4} + \cdots\right], \\ &= r_0 + \frac{\rho^2}{2r_0} - \frac{\rho^4}{8r_0^3} + \cdots . \end{aligned} \tag{B.6}$$

The Fraunhofer approximation to the range is found by discarding all the non-linear terms in the expansion:

$$r \approx r_0 + \frac{(\mathbf{x}_0 - \mathbf{x}_s)\cdot\mathbf{x}'}{r_0}. \tag{B.7}$$

This approximation simplifies the analysis of many situations involving radiating energy. For example, in the simulator scattering model presented in Section 4.1, the application of the Fraunhofer approximation allows the replacement of a surface integral with a Fourier transform.

B.1 GEOMETRIC INTERPRETATION

Using the definition of the dot product, $\mathbf{a}\cdot\mathbf{b} = |\mathbf{a}||\mathbf{b}|\cos\theta$, where θ is the angle between the vectors, the approximated range of (B.7) can be written as

$$\begin{aligned} r &\approx r_0 + \frac{|\mathbf{x}_0 - \mathbf{x}_s||\mathbf{x}'|\cos\theta}{r_0}, \\ &\approx r_0 + \left|\mathbf{x}'\right|\cos\theta. \end{aligned} \tag{B.8}$$

This can then be viewed as a projection of the offset vector $\mathbf{x}'$ onto the second measurement vector $\mathbf{x}_0$. This is illustrated for a simple two-dimensional case in Figure B.1. As drawn, the range is approximated by the horizontal component only, i.e., the vertical component is discarded. This has the effect of replacing the spherical wavefronts of the radiated energy with planar wavefronts. As shown in Figure B.2, this is a reasonable approximation for a portion of the wavefront, and this portion gets larger as the range from the source increases.

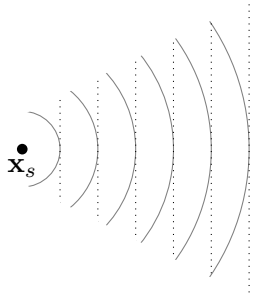


Figure B.2: The Fraunhofer approximation has the effect of replacing the true spherical wavefronts of the radiated energy (drawn here as solid lines) with planar waves (the dotted lines) over a small portion of the wavefront normal to some desired direction.

B.2 REGION OF VALIDITY

The region in which the Fraunhofer approximation is considered valid is known as the Fraunhofer region. This consists of all space beyond the Rayleigh distance R which is the distance at which the error in the approximation is at most one-sixteenth of the wavelength.[2] This scenario is shown in Figure B.3. There is a region of length L where it is desired to use the Fraunhofer approximation.[3] The range to the centre of the region is set as the Rayleigh distance R, and the maximum error $R + \lambda/16$ will occur when approximating the range to a point at the edge of the region. Applying the Pythagorean theorem to the right-angled triangle this creates gives

$$\left(R + \frac{\lambda}{16}\right)^2 = R^2 + \left(\frac{L}{2}\right)^2. \tag{B.9}$$

This can be rearranged to give the Rayleigh distance as

$$R = \frac{2L^2}{\lambda} - \frac{\lambda}{32}. \tag{B.10}$$

The constant term is commonly dropped, and the distance can alternatively be expressed in terms of the frequency f and speed of propagation c of the wave:

$$\begin{aligned} R &= \frac{2L^2}{\lambda}, \\ &= \frac{2fL^2}{c}. \end{aligned} \tag{B.11}$$

For ranges shorter than this, the Fresnel approximation can be used; this also retains the quadratic terms in the binomial expansion of (B.6). This effectively replaces the spherical wavefronts with quadratic versions, and is valid to within a few wavelengths

[2] In optics, the Rayleigh distance is defined for a maximum error of one-quarter of a wavelength. Antenna applications generally use the version given here.

[3] For a facet, it is desired to use the approximation over the entire facet. In this case, L should be the largest extent of the facet.

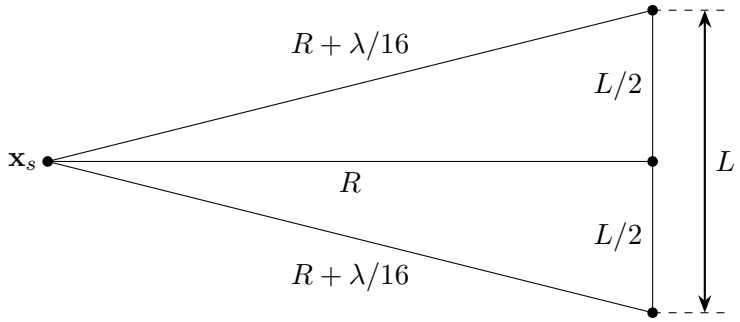


Figure B.3: The derivation of the Rayleigh distance. The Fraunhofer approximation is used for the range from the source $\mathbf{x}_s$ to the points in a region of length L. The Rayleigh distance R is the distance at which the true range differs from the approximated version by no more than one-sixteenth of the wavelength.

of the source.

B.3 EFFECT ON REPEAT-PASS COHERENCE

At first glance, an error of $\lambda/16$ (equivalent to a phase shift of 22.5° or 0.393 radians) may seem large enough to have a significant effect on the repeat-pass coherence of the simulated data. However, it should be borne in mind that this is the maximum difference between the approximated phase and the true phase in one pass. If the same scene is re-simulated with an identical track then the same approximation errors will be introduced, and the interferometric phase between the two images will be zero (i.e., they will be perfectly coherent). Even for realistic track separations, the error is not significant. For example, consider a scenario with a wavelength $\lambda = 5\,\text{cm}$ where it is desired to apply the Fraunhofer approximation over a length $L = 50\,\text{cm}$. The corresponding Rayleigh distance is $R = 10\,\text{m}$, and the maximum error for a point imaged at this distance is $\lambda/16 = 3.125\,\text{mm}$. Now assume that a repeat pass is performed at a distance of 11 m from the region of interest, i.e., a 1 m difference from the primary pass. The maximum error resulting from applying the Fraunhofer approximation over the length L in this pass is 2.841 mm. The (maximum) difference in the approximation errors between these two passes is 0.284 mm, or approximately $\lambda/176$.

It is also worth noting that this is a worst-case error. The approximated ranges to all but the end points of the length L have an error below the $\lambda/16$ limit. In an actual simulation, the value of L will be chosen based on the minimum range between the sonar and the scene being simulated. The majority of the scene will be more distant from the sonar and so have a smaller approximation error. If the error is deemed to be too large for a given application, then it can be reduced by decreasing the value of L; however, this will increase the computational workload.

Appendix C

STATISTICAL CORRELATION

The correlation between two signals as a function of the delay or lag between them in defined in Section 3.7. The concept of correlation also exists in statistics for both random variables and samples of random variables. For completeness, these definitions are given in Section C.1 and Section C.2. Some caution is required for the interpretation of correlation values as they only give details of the linear relationship between the two variables. This is detailed in Section C.3.

C.1 ENSEMBLE CORRELATION

For two jointly-distributed complex random variables X and Y, the covariance is defined as [Clapham and Nicholson 2009]

$$\text{cov}(X, Y) = \text{E}[(X - \text{E}[X])(Y - \text{E}[Y])^*]. \tag{C.1}$$

This is a measure of how similar the variations in the variables are. If large values of X correspond to large values of Y, and the small values also correspond, then the covariance will be positive. Similarly, the covariance will be negative if the large values of X correspond to the small values of Y and vice-versa. If there is no such correspondence, i.e., if the variables are independent, then the covariance will be zero [Bulmer 2012]. Note that the covariance of a random variable with itself is equivalent to its variance, i.e.,

$$\text{cov}(X, X) = \text{Var}[X]. \tag{C.2}$$

It is difficult to infer the level of correspondence between the variables solely from their covariance. For example, suppose both X and Y are multiplied by some real constant a. Applying the linearity property of expectation to (C.1), the covariance can then be

shown to be

$$\begin{aligned}\operatorname{cov}(aX, aY) &= \mathrm{E}[(aX - \mathrm{E}[aX])(aY - \mathrm{E}[aY])^*], \\ &= \mathrm{E}[(aX - a\,\mathrm{E}[X])(aY - a\,\mathrm{E}[Y])^*], \\ &= \mathrm{E}\big[a^2(X - \mathrm{E}[X])(Y - \mathrm{E}[Y])^*\big], \\ &= a^2 \operatorname{cov}(X, Y).\end{aligned} \tag{C.3}$$

Although the same relative relationship exists between the multiplied variables as between the original variables, the value of the covariance has increased. To avoid this, the covariance can be normalised by the product of the standard deviations of the variables [Pearson 1895], yielding the Pearson correlation coefficient[1]

$$\operatorname{corr}(X, Y) = \frac{\operatorname{cov}(X, Y)}{\sigma_X \sigma_Y}. \tag{C.4}$$

The normalisation ensures the correlation coefficient takes some value in the closed interval $[-1, 1]$. If the two variables are identical then $\sigma_X = \sigma_Y$, from (C.2) the covariance is σ_X^2, and therefore the correlation coefficient is 1. Similarly, if the variables have a perfect negative relationship (when one takes large values the other takes small and vice-versa), the correlation coefficient will be -1 (sometimes referred to as anticorrelation). Values in between these limits indicate the degree of linear dependence between the variables, with a correlation coefficient of zero corresponding to there being no linear dependence (the variables are uncorrelated).

C.2 SAMPLE CORRELATION

For two vectors $\mathbf{x}$ and $\mathbf{y}$, each containing N observed values of X and Y, the sample covariance can be computed as

$$\operatorname{cov}(\mathbf{x}, \mathbf{y}) = \frac{1}{N-1} \sum_{i=1}^{N} (x_i - \overline{\mathbf{x}})(y_i - \overline{\mathbf{y}})^*, \tag{C.5}$$

where $\overline{\mathbf{x}}$ and $\overline{\mathbf{y}}$ are the sample means of $\mathbf{x}$ and $\mathbf{y}$ respectively. As the sample means are also calculated from the observation vectors, there is one less degree of freedom when calculating the sample covariance, and so $N - 1$ is used in the denominator as using N would yield a biased estimate [Johnson 2007].[2] If the population mean is known, then $\mathrm{E}[X]$ and $\mathrm{E}[Y]$ replace the sample means in (C.5) and the denominator becomes N. Dividing (C.5) by the expressions for the sample standard deviations of $\mathbf{x}$ and $\mathbf{y}$ gives

[1] There are other correlation coefficients, for example Spearman's rank correlation coefficient. The Pearson correlation coefficient is the most widely used, and is often referred to (as this thesis does) as the correlation coefficient.

[2] This is known as Bessel's correction. Although it corrects the bias, it may increase the mean squared error of the estimate.

the sample correlation coefficient:

$$\operatorname{corr}(\mathbf{x},\mathbf{y}) = \frac{\sum_{i=1}^{N}(x_i-\overline{\mathbf{x}})(y_i-\overline{\mathbf{y}})^*}{\sqrt{\sum_{i=1}^{N}|x_i-\overline{\mathbf{x}}|^2}\sqrt{\sum_{i=1}^{N}|y_i-\overline{\mathbf{y}}|^2}}. \tag{C.6}$$

As with the population correlation coefficient, the sample correlation coefficient lies on the interval $[-1,1]$ and is a measure of the linear dependence between the two observation vectors. Comparing this to the formula for calculating the correlation of a sampled signal given in (3.40) shows that the only difference is the presence of the lag τ in the signal processing version.

C.3 LIMITATIONS

Non-linear relationships are not guaranteed to be detected by the correlation. For example, let X be uniformly distributed on the interval $[-1,1]$. Squaring this generates a new random variable, $Y = X^2$, which is by definition dependent on X. However, calculating the population covariance as per (C.1) gives

$$\begin{aligned}
\operatorname{cov}(X,Y) &= \mathrm{E}\big[(X-\mathrm{E}[X])\big(X^2-\mathrm{E}\big[X^2\big]\big)\big],\\
&= \mathrm{E}\big[X^3 - X\,\mathrm{E}\big[X^2\big]\big],\\
&= \mathrm{E}\big[X^3\big] - \mathrm{E}[X]\,\mathrm{E}\big[X^2\big],\\
&= 0 - 0\times\mathrm{E}\big[X^2\big],\\
&= 0.
\end{aligned} \tag{C.7}$$

As a result, the correlation coefficient is also zero. Although Y is dependent on X, it is not a linear dependency and hence the correlation coefficient does not detect it.

A corollary to this is that the correlation coefficient does not say anything about the relationship between the variables, but rather to what extent it can be approximated as a linear relationship. A well known example of this is Anscombe's quartet [Anscombe 1973]. Plotted in Figure C.1, this consists of four datasets of (x,y) pairs. Although the graphs show they are markedly different, the means of x (9) and y (7.5), their sample variances (11 and 4.125 respectively), the correlation coefficient between x and y (0.816), and the linear regression line ($y = 3+0.5x$) are identical for all four datasets. This illustrates that, while useful, summary statistics should not be the sole piece of information relied on when analysing data.

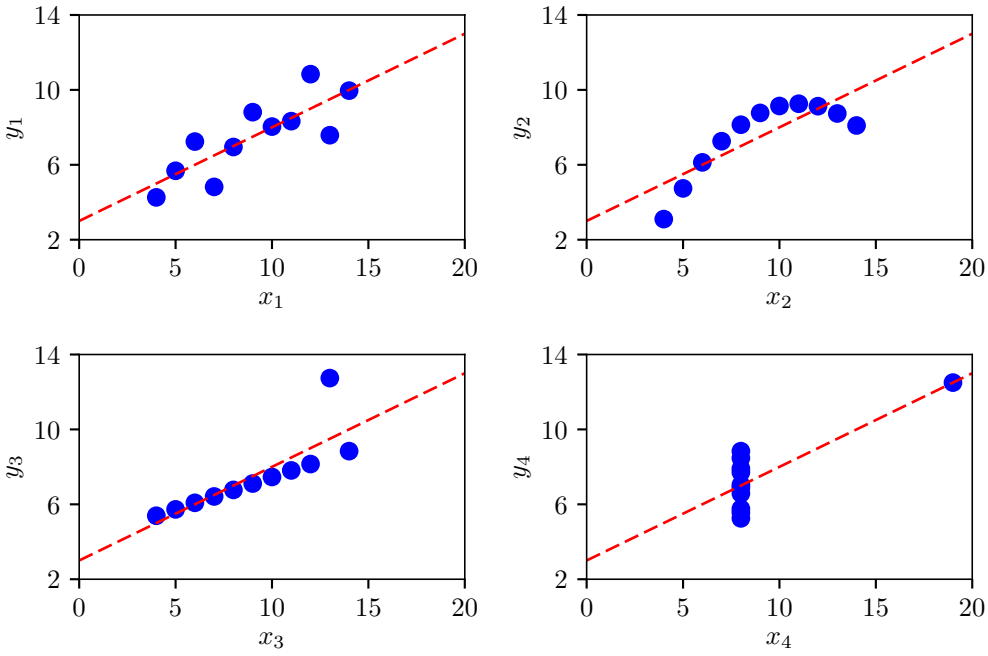


Figure C.1: Anscombe's quartet, four datasets of (x, y) pairs which have identical means, variances, correlation coefficients and linear regression lines (shown in red) yet are markedly different. The quartet was constructed by Anscombe [1973] to highlight that analysis of data should not rely solely on summary statistics.

REFERENCES

Andrew, R. K., Howe, B. M., Mercer, J. A., and Dzieciuch, M. A. (2002). "Ocean ambient sound: Comparing the 1960s with the 1990s for a receiver off the California coast". In: *Acoustics Research Letters Online* 3.2, pp. 65–70.

Anscombe, F. J. (1973). "Graphs in Statistical Analysis". In: *The American Statistician* 27.1, pp. 17–21.

Au, W. W. L. (1993). *The Sonar of Dolphins*. Springer.

Au, W. W. L. and Banks, K. (1998). "The acoustics of the snapping shrimp *Synalpheus parneomeris* in Kaneohe Bay". In: *Journal of the Acoustical Society of America* 103.1, pp. 41–47.

Barclay, P. J. (2006). "Interferometric Synthetic Aperture Sonar Design and Performance". PhD thesis. University of Canterbury.

Bay, H., Tuytelaars, T., and Van Gool, L. (2006). "SURF: Speeded up robust features". In: *European Conference on Computer Vision*, pp. 404–417.

Bellec, R., Legris, M., Khenchaf, A., Amate, M., and Hetet, A. (2005). "Repeat-track SAS interferometry: feasibility study". In: *Proceedings of MTS/IEEE OCEANS 2005*. Vol. 1, pp. 748–754.

Benoît, E. (1924). "Note sur une méthode de résolution des èquations normales provenant de l'application de la méthode des moindres carrés à un système d'équations linéaires en nombre inférieur à celui des inconnues (Procédé du Commandant Cholesky) [Note on a method for solving the normal equations resulting from the application of the least-squares method to a system of linear equations in numbers lower than that of the unknowns (Commandant Cholesky's method)]". In: *Bulletin Géodésique* 2, pp. 67–77.

Birkhoff, G. D. (1931). "Proof of the Ergodic Theorem". In: *Proceedings of the National Academy of Sciences* 17.12, pp. 656–660.

Bonnett, B. and Hayes, M. P. (2014). "Data-driven Image Registration for Coherent Change Detection of Synthetic Aperture Sonar Imagery". In: *Proceedings of the 29th*

International Conference on Image and Vision Computing New Zealand. IVCNZ '14. Hamilton, New Zealand: ACM, pp. 196–201.

Bonnett, B. and Hayes, M. P. (2016). "Simulation of temporal coherence loss for repeat-pass synthetic aperture sonar". In: *2016 International Conference on Image and Vision Computing New Zealand*. IVCNZ '16. Palmerston North, New Zealand.

Bonnett, B., Hayes, M. P., and Hunter, A. (2013). "Registration of images from a hull mounted, low frequency synthetic aperture sonar". In: *28th International Conference on Image and Vision Computing New Zealand*. IVCNZ 2013. Wellington, New Zealand, pp. 142–147.

Born, M. and Wolf, E. (1999). *Principles of Optics*. 6th ed. Cambridge University Press.

Boucher, R. E. and Hassab, J. C. (1981). "Analysis of discrete implementation of generalized cross correlator". In: *IEEE Transactions on Acoustics, Speech, and Signal Processing* 29.3, pp. 609–611.

Bracewell, R. N. (2000). *The Fourier Transform and Its Applications*. 3rd ed. McGraw-Hill.

Bray, J. R. and Curtis, J. T. (1957). "An ordination of the upland forest communities of southern Wisconsin". In: *Ecological Monographs* 27, pp. 325–349.

Bruzzone, L. and Prieto, D. F. (2000). "Automatic Analysis of the Difference Image for Unsupervised Change Detection". In: *IEEE Transactions on Geoscience and Remote Sensing* 38.3, pp. 1171–1182.

Bryan, M. L. and Clark, J. (1984). "Potentials for change detection using Seasat synthetic aperture radar data". In: *Remote Sensing of Environment* 16.2, pp. 107–124.

Bulmer, M. G. (2012). *Principles of Statistics*. Dover Books on Mathematics. Dover Publications.

Byrd, R. H., Lu, P., and Nocedal, J. (1995). "A Limited Memory Algorithm for Bound Constrained Optimization". In: *SIAM Journal on Scientific and Statistical Computing* 16.5, pp. 1190–1208.

Cafforio, C., Pratti, C., and Rocca, F. (1991). "SAR data focusing using seismic migration techniques". In: *IEEE Transactions on Aerospace and Electronic Systems* 27.2, pp. 194–207.

Callow, H. J. (2003). "Signal Processing for Synthetic Aperture Sonar Image Enhancement". PhD thesis. University of Canterbury.

Campbell, J. B. (2008). "Origins of Aerial Photographic Interpretation, U.S. Army, 1916 to 1918". In: *Photogrammetric Engineering and Remote Sensing* 74.1, pp. 77–93.

Cantelli, F. P. (1933). "Sulla determinazione empirica delle leggi di probabilità [On the empirical determination of the probability laws]". In: *Giornale Dell'Istituto Italiano Degli Attuari* 4, pp. 421–424.

Caprais, P. and Guyonic, S. (1997). "Squint and forward looking synthetic aperture sonar". In: *OCEANS '97. MTS/IEEE Conference Proceedings*. Vol. 2, pp. 809–814.

Carter, G. C. (1987). "Coherence and Time Delay Estimation". In: *Proceedings of the IEEE* 75.2, pp. 236–255.

Chi, H., Sun, G., and Ling, F. (2009). "Urban dynamic change detection in southeastern China based on interferometric SAR". In: *2009 IEEE International Geoscience and Remote Sensing Symposium*. Vol. 3, pp. 432–435.

Cho, Z. H., Jones, J. P., and Singh, M. (1993). *Foundations of Medical Imaging*. Wiley Interscience.

Cholesky, A.-L. (1910). "Sur la résolution numérique des systèmes d'équations linéaires [On the numerical solution of systems of linear equations]". Originally unpublished and found with the authors personal papers, digitised by the École Polytechnique in 2005.

Clapham, C. and Nicholson, J. (2009). *The Concise Oxford Dictionary of Mathematics*. 4th ed. Oxford Quick Reference. Oxford University Press.

Coiras, E., Groen, J., Williams, D., Evans, B., and Pinto, M. (2008). "Automatic change detection for the monitoring of cluttered underwater areas". In: *Proceedings of the 1st International Conference on Waterside Security*.

Coppin, P., Jonckheere, I., Nackaerts, K., Muys, B., and Lambin, E. (2004). "Digital change detection methods in ecosystem monitoring: a review". In: *International Journal of Remote Sensing* 25.9, pp. 1565–1596.

Corr, D. G. (1997). "Coherent change detection for urban development monitoring". In: *IEE Colloquium on Radar Interferometry (Digest No: 1997/153)*, pp. 6/1–6/6.

Corr, D. G. and Rodrigues, A. F. (1998). "Coherent change detection of vehicle movements". In: *Proceedings of the 1998 Geoscience and Remote Sensing Symposium*. Vol. 5, pp. 2451–2453.

Corr, D. G. and Rodrigues, A. F. (1999). "Change detection using interferometric SAR data". In: *Proceedings of the SPIE* 3869, pp. 127–138.

Corr, D. G., Whitehouse, S. W., Mott, D. H., and Baldwin, J. F. (1996). "Automatic change detection in spaceborne SAR imagery". In: *Proceedings of the SPIE* 2757, pp. 58–69.

Craig, C. C. (1936). "On the Frequency Function of xy". In: *The Annals of Mathematical Statistics* 7.1, pp. 1–15.

Crocker, M. J. (1998). *Handbook of Acoustics*. New York: Wiley Interscience.

Croswell, K. (1997). *Planet Quest: The Epic Discovery of Alien Solar Systems*. Free Press.

Crow, F. C. (1984). "Summed-area Tables for Texture Mapping". In: *SIGGRAPH Computer Graphics* 18.3, pp. 207–212.

Curlander, J. C. and McDonough, R. N. (1991). *Synthetic Aperture Radar: Systems and Signal Processing*. New York: John Wiley and sons.

Dainty, J. C., ed. (1975). *Laser Speckle and Related Phenomena*. Springer-Verlag.

Dainty, J. C. (1976). "The Statistics of Speckle Patterns". In: *Progress in Optics, vol. XIV*. Ed. by E. Wolf. North-Holland Publishing Company. Chap. 1, pp. 3–46.

De Paulis, R., Prati, C., Scirpoli, S., Rocca, F., Tesei, A., Sletner, P. A., Biagini, S., Guerrini, P., Gasparoni, F., Carmisciano, C., and Locritani, M. (2011). "SAS multipass interferometry for monitoring seabed deformation using a high-frequency imaging sonar". In: *OCEANS 2011 IEEE – Spain*.

Dekker, R. J. (2003). "Texture analysis and classification of ERS SAR images for map updating of urban areas in The Netherlands". In: *IEEE Transactions on Geoscience and Remote Sensing* 41.9, pp. 1950–1958.

Dillon, J. and Myers, V. (2014). "Coherence estimation for repeat-pass interferometry". In: *OCEANS 2014*. St. John's.

Engeset, R. V., Kohler, J., Melvold, K., and Lundén, B. (2002). "Change detection and monitoring of glacier mass balance and facies using ERS SAR winter images over Svalbard". In: *International Journal of Remote Sensing* 23.10, pp. 2023–2050.

Ennos, A. E. (1996). "Laser speckle experiments for students". In: *Physics Education* 31.3, pp. 138–142.

Exner, K. (1878). "Über die Fraunhofer'schen Ringe, die Quetelet'schen Streifen und verwandte Erscheinungen [Fraunhofer's rings, Quetelet's stripes, and related phenomena]". In: *Annalen der Physik* 240.8, pp. 525–550.

Exner, K. (1880). "Über die Newton'schen Staubringe [About the Newtonian dust rings]". In: *Annalen der Physik* 245.2, pp. 239–260.

Facchinetti, S. (2009). "A procedure to find exact critical values of Kolmogorov-Smirnov test". In: *Statistica Applicata — Italian Journal of Applied Statistics* 21.3–4, pp. 337–359.

Fahy, F. (1998). *Fundamentals of noise and vibration*. Taylor & Francis.

Ferrand, J. and Mandelert, N. (2012). "Change detection for MCM survey mission". In: *Proceedings of the International Conference on Underwater Remote Sensing*. Brest, France, pp. 193–206.

Fily, M. and Rothrock, D. A. (1987). "Sea Ice Tracking by Nested Correlations". In: *IEEE Transactions on Geoscience and Remote Sensing* GE-25.5, pp. 570–580.

Fisher, F. H. and Simmons, V. P. (1977). "Sound absorption in sea water". In: *Journal of the Acoustical Society of America* 62, pp. 558–564.

Fortune, S. A. (2005). "Phase Error Estimation for Synthetic Aperture Imagery". PhD thesis. University of Canterbury.

Fried, D. L. (1966). "Optical Resolution Through a Randomly Inhomogeneous Medium for Very Long and Very Short Exposures". In: *Journal of the Optical Society of America* 56.10, pp. 1372–1379.

Gallager, R. G. (2008). "Circularly-symmetric Gaussian random vectors". Unpublished additions to the book "Principles of Digital Communication", available at `http://www.rle.mit.edu/rgallager/documents/CircSymGauss.pdf`.

Gamba, P., Dell'Acqua, F., and Lisini, G. (2006). "Change Detection of Multitemporal SAR Data in Urban Areas Combining Feature-Based and Pixel-Based Techniques". In: *IEEE Transactions on Geoscience and Remote Sensing* 44.10, pp. 2820–2827.

Gendron, M. L. and Lohrenz, M. C. (2007). "The Automated Change Detection and Classification Real-time (ACDC-RT) System". In: *OCEANS 2007 - Europe*.

Gendron, M. L., Lohrenz, M., Layne, G., and Ioup, J. (2004). "Automatic Change Detection and Classification (ACDC) System". In: *Proceedings of the Sixth International Symposium on Technology and the Mine Problem*.

Gendron, M., Lohrenz, M., and Dubberley, J. (2009). "Automated change detection using Synthetic Aperture Sonar imagery". In: *OCEANS 2009, MTS/IEEE Biloxi - Marine Technology for Our Future: Global and Local Challenges*. Biloxi, MS.

Gerónimo, D., López, A. M., Sappa, A. D., and Graf, T. (2010). "Survey of Pedestrian Detection for Advanced Driver Assistance Systems". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 32.7, pp. 1239–1258.

Gierull, C. H. and Sikaneta, I. C. (2002). "Estimating the Effective Number of Looks in Interferometric SAR Data". In: *IEEE Transactions on Geoscience and Remote Sensing* 40.8, pp. 1733–1742.

Glivenko, V. I. (1933). "Sulla determinazione empirica delle leggi di probabilità [On the empirical determination of the probability laws]". In: *Giornale Dell'Istituto Italiano Degli Attuari* 4, pp. 92–99.

G-Michael, T., Marchand, B., Tucker, J. D., Marston, T. M., Sternlicht, D. D., and Azimi-Sadjadi, M. R. (2016). "Image-Based Automated Change Detection for Synthetic Aperture Sonar by Multistage Coregistration and Canonical Correlation Analysis". In: *IEEE Journal of Oceanic Engineering* 41.3, pp. 592–612.

G-Michael, T., Marchand, B., Tucker, J. D., Sternlicht, D. D., Marston, T. M., and Azimi-Sadjadi, M. R. (2014). "Automated change detection for synthetic aperture sonar". In: *Proceedings of the SPIE* 9072.

G-Michael, T. and Tucker, J. D. (2010). "Canonical correlation analysis for coherent change detection in synthetic aperture sonar imagery". In: *International Conference on Synthetic Aperture Sonar and Synthetic Aperture Radar 2010*. Proceedings of the Institute of Acoustics Volume 32 Pt.4. Lerici, Italy: Curran Associates, Inc, pp. 117–122.

G-Michael, T., Tucker, J. D., and Roberts, R. G. (2016). "Statistically normalized coherent change detection for synthetic aperture sonar imagery". In: *Proceedings of the SPIE* 9823, 98231T-1 –98231T-6.

Goldberg, D. (1991). "What Every Computer Scientist Should Know About Floating Point Arithmetic". In: *ACM Computing Surveys* 23.1, pp. 5–48.

Goodman, J. W. (1975). "Statistical properties of laser speckle patterns". In: *Laser speckle and related phenomena*. Ed. by J. C. Dainty. Vol. 9. Topics in Applied Physics. Berlin: Springer-Verlag. Chap. 1, pp. 9–75.

Goodman, J. W. (1976). "Some fundamental properties of speckle". In: *Journal of the Optical Society of America* 66.11, pp. 1145–1150.

Goodman, J. W. (1986). "A random walk through the the field of speckle". In: *Optical Engineering* 25.5, pp. 610–612.

Goodman, N. R. (1963). "Statistical Analysis Based on a Certain Multivariate Complex Gaussian Distribution (An Introduction)". In: *The Annals of Mathematical Statistics* 34.1, pp. 152–177.

Guarnieri, A. M. and Prati, C. (1997). "SAR interferometry: a 'Quick and dirty' coherence estimator for data browsing". In: *IEEE Transactions on Geoscience and Remote Sensing* 35.3, pp. 660–669.

Hamilton, E. L. (1980). "Geoacoustic modelling of the sea floor". In: *Journal of the Acoustical Society of America* 68, pp. 1313–1340.

Hamilton, E. L. and Bachman, R. T. (1982). "Sound velocity and related properties in marine sediments". In: *Journal of the Acoustical Society of America* 72, pp. 1891–1904.

Hansen, R. E., Lyons, A. P., Sæbø, T. O., Callow, H. J., and Cook, D. A. (2015). "The Effect of Internal Wave-Related Features on Synthetic Aperture Sonar". In: *IEEE Journal of Oceanic Engineering* 40.3, pp. 621–631.

Hanssen, R. and Bamler, R. (1999). "Evaluation of interpolation kernels for SAR interferometry". In: *IEEE Transactions on Geoscience and Remote Sensing* 37.1, pp. 318–321.

Hariharan, P. (1972). "Speckle Patterns: A Historical Retrospect". In: *Optica Acta* 19.9, pp. 791–793.

Hawkins, D. W. and Gough, P. T. (1995). "Recent sea trials of a synthetic aperture sonar". In: *Proceedings of the Institute of Acoustics* 17.8.

Hawkins, D. W. (1996). "Synthetic Aperture Imaging Algorithms: with application to wide bandwidth sonar". PhD thesis. University of Canterbury.

Hayes, M. P. and Gough, P. T. (1992). "Broadband synthetic aperture sonar". In: *IEEE Journal of Oceanic Engineering* 17.1, pp. 80–94.

Hayes, M. P. and Gough, P. T. (2008). "Interferometric synthetic aperture processing: a comparison of sonar and radar". In: *Acoustics'08 Paris, the 2nd joint conference of the Acoustical Society of America (ASA) and the European Acoustics Association (EAA)*. Paris, pp. 6303–6308.

Hibler III, W. D., Ackley, S. F., Crowder, W. K., McKim, H. L., and Anderson, D. M. (1974). "Analysis of shear zone ice deformation in the Beaufort Sea using satellite imagery". In: *The coast and shelf of the Beaufort Sea: proceedings of a Symposium on Beaufort Sea Coast and Shelf Research*. Ed. by J. C. Reed and J. E. Sater. Artic Institute of North America.

Hibler III, W. D., Tucker, W. B., and Weeks, W. F. (1975). *Techniques for Studying Sea Ice Drift and Deformation at Sites Far from Land Using LANDSAT Imagery*. Tech. rep. ADA041579. Hanover, NH, USA: Cold Regions Research and Engineering Lab.

Higham, N. J. (1993). "The Accuracy of Floating Point Summation". In: *SIAM Journal on Scientific Computing* 14.4, pp. 783–799.

Hotelling, H. (1936). "Relations between two sets of variates". In: *Biometrika* 28.3–4, pp. 321–377.

Huang, T. S., Burnett, J. W., and Deczky, A. G. (1975). "The Importance of Phase in Image Processing Filters". In: *IEEE Transactions on Acoustics, Speech, and Signal Processing* 23.6, pp. 529–542.

Hughes, J. F., Dam, A. van, McGuire, M., Sklar, D. F., Foley, J. D., Feiner, S. K., and Akeley, K. (2013). *Computer graphics: principles and practice*. 3rd ed. Boston: Addison-Wesley Professional.

Hunt, F. V. (1978). *Origins in Acoustics: the science of sound from antiquity to the age of Newton*. Yale University Press.

Hunter, A. J., Dugelay, S., and Fox, W. L. J. (2016). "Repeat-Pass Synthetic Aperture Sonar Micronavigation Using Redundant Phase Center Arrays". In: *IEEE Journal of Oceanic Engineering* 41.4, pp. 820–830.

Hunter, A. J. (2006). "Underwater Acoustic Modelling for Synthetic Aperture Sonar". PhD thesis. University of Canterbury.

IEEE (2008). "IEEE Standard for Floating-Point Arithmetic". In: *IEEE Std 754-2008*.

Jackson, D. R., Richardson, M. D., Williams, K. L., Lyons, A. P., Jones, C. D., Briggs, K. B., and Tang, D. (2009). "Acoustic Observation of the Time Dependence of the Roughness of Sandy Seafloors". In: *IEEE Journal of Oceanic Engineering* 34.4, pp. 407–422.

Jackson, D. R., Williams, K. L., and Briggs, K. B. (1996). "High-frequency acoustic observations of benthic spatial and temporal variability". In: *Geo-Marine Letters* 16.3, pp. 212–218.

Jackson, D. and Richardson, M. (2007). *High-Frequency Seafloor Acoustics*. The Underwater Acoustics Series. New York: Springer.

Jin, G. and Tang, D. (1996). "Uncertainties of differential phase estimation associated with interferometric sonars". In: *IEEE Journal of Oceanic Engineering* 21, pp. 53–63.

Johnson, R. A. (2007). *Applied Multivariate Statistical Analysis*. 6th ed. Pearson Prentice Hall.

Jonsson, M., Pihl, J., and Aklint, M. (2005). "Imaging of buried objects by low frequency SAS". In: *OCEANS 2005 Europe*. Vol. 1, pp. 669–673.

Just, D. and Bamler, R. (1994). "Phase statistics of interferograms with applications to synthetic aperture radar". In: *Applied Optics* 33.20, pp. 4361–4368.

Kan, W. Y., Krogmeier, J. V., and Doerschuk, P. C. (1996). "Model-based vehicle tracking from image sequences with an application to road surveillance". In: *Optical Engineering* 35.6, pp. 1723–1729.

Kinsey, J. C., Eustice, R. M., and Whitcomb, L. L. (2006). "A Survey of Underwater Vehicle Navigation: Recent Advances and New Challenges". In: *IFAC Conference of Manoeuvering and Control of Marine Craft*. Vol. 88.

Kirchhoff, G. R. (1882). "Zur Theorie der Lichtstrahlen [The theory of light beams]". In: *Annalen der Physik* 254.4, pp. 663–695.

Kirlin, R. L., Lu, W.-S., Hedstrom, B., and Leung, C.-M. (1993). "Towfish Orientation and Position Estimation". In: *IEEE Journal of Oceanic Engineering* 18.3, pp. 319–326.

Kolmogorov, A. N. (1933). "Sulla determinazione empirica di una legge di distribuzione [On the empirical determination of a distribution law]". In: *Giornale dell'Istituto Italiano degli Attuari* 4, pp. 83–91.

Kuch, B., Buttazzo, G., Azzopardi, E., Sayer, M., and Sieber, A. (2012). "GPS diving computer for underwater tracking and mapping". In: *Underwater Technology* 30.4, pp. 189–194.

Kullback, S. and Leibler, R. A. (1951). "On information and sufficiency". In: *Annals of Mathematical Statistics* 22.1, pp. 79–86.

Kuttikkad, S. and Chellappa, R. (2000). "Statistical modeling and analysis of high-resolution Synthetic Aperture Radar images". In: *Statistics and Computing* 10.2, pp. 133–145.

Landis, E. N., Nagy, E. W., Keane, D. T., and Nagy, G. (1999). "Technique to measure 3D work-of-fracture of concrete in compression". In: *Journal of Engineering Mechanics* 125.6, pp. 599–605.

Łapucha, D., Jong, K. de, Liu, X., Melgard, T., Oerpen, O., and Vigen, E. (2011). "Recent advances in Wide Area Real-Time Precise Positioning". In: *Navigational Systems and Simulators: Marine Navigation and Safety of Sea Transportation*. Ed. by A. Weintrit. CRC Press. Chap. 6, pp. 49–54.

Leberl, F., Bryan, M. L., Elachi, C., Farr, T., and Campbell, W. (1979). "Mapping of sea ice and measurement of its drift using aircraft synthetic aperture radar images". In: *Journal of Geophysical Research: Oceans* 84.C4, pp. 1827–1835.

Leberl, F., Raggam, J., Elachi, C., and Campbell, W. J. (1983). "Sea Ice Motion Measurements From SEASAT SAR Images". In: *Journal of Geophysical Research* 88.C3, pp. 1915–1928.

Lee, J.-S., Hoppel, K. W., Mango, S. A., and Miller, A. R. (1994). "Intensity and phase statistics of multilook polarimetric and interferometric SAR imagery". In: *IEEE Transactions on Geoscience and Remote Sensing* 32.5, pp. 1017–1028.

Leo, W. R. (1994). *Techniques for Nuclear and Particle Physics Experiments: A How-To Approach*. Springer.

Leonard, J. L., Bennett, A. A., Smith, C. M., Jacob, H., and Feder, S. (1998). *Autonomous Underwater Vehicle Navigation*. Tech. rep. MIT Marine Robotics Laboratory.

Lewis, J. P. (1995). "Fast template matching". In: *Vision Interface 95*. (Quebec City, Canada, May 15–19, 1995). Canadian Image Processing and Pattern Recognition Society, pp. 120–123.

Li, Q. (2012). *Digital Sonar Design in Underwater Acoustics: Principles and Applications*. Springer Science & Business Media.

Liu, J. G., Black, A., Lee, H., Hanaizumi, H., and Moore, J. M. (2001). "Land surface change detection in a desert area in Algeria using multi-temporal ERS SAR coherence images". In: *International Journal of Remote Sensing* 22.13, pp. 2463–2477.

Løkberg, O. J. (1980). "Electronic speckle pattern interferometry". In: *Physics in Technology* 11.1, pp. 16–22.

Lowe, D. G. (1999). "Object recognition from local scale-invariant features". In: *Proceedings of the Seventh IEEE International Conference on Computer Vision*. Vol. 2, pp. 1150–1157.

Lu, D., Mausel, P., Brondízio, E., and Moran, E. (2004). "Change detection techniques". In: *International Journal of Remote Sensing* 25.12, pp. 2365–2407.

Lunetta, R. S. and Elvidge, C. (1998). *Remote sensing change detection: environmental monitoring methods and applications*. Ann Arbor Press.

Lyons, A. P. and Abraham, D. A. (1999). "Statistical characterization of high-frequency shallow-water seafloor backscatter". In: *The Journal of the Acoustical Society of America* 106, pp. 1307–1315.

Lyons, A. P., Abraham, D. A., and Johnson, S. F. (2010). "Modeling the effect of seafloor ripples on synthetic aperture sonar speckle statistics". In: *IEEE Journal of Oceanic Engineering* 35.2, pp. 242–249.

Lyons, A. P. and Brown, D. C. (2013). "The Impact of the Temporal Variability of Seafloor Roughness on Synthetic Aperture Sonar Repeat-Pass Interferometry". In: *IEEE Journal of Oceanic Engineering* 38.1, pp. 91–97.

Madgwick, S. O. H. (2010). *An efficient orientation filter for inertial and inertial/magnetic sensor arrays*. Tech. rep. University of Bristol.

Mandelbrot, B. B. (1982). *The Fractal Geometry of Nature*. New York: W. H. Freeman and Company.

Manolakis, D. G., Ingle, V. K., and Kogon, S. M. (2005). *Statistical and Adaptive Signal Processing*. Artech House.

Marins, J. L., Yun, X., Bachmann, E. R., and Zyda, M. J. (2001). "An extended Kalman filter for quaternion-based orientation estimation using MARG sensors". In: *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*. Vol. 4, pp. 2003–2011.

Mas, J.-F. (1999). "Monitoring land-cover changes: A comparison of change detection techniques". In: *International Journal of Remote Sensing* 20.1, pp. 139–152.

Massey, F. J. (1951). "The Kolmogorov-Smirnov Test for Goodness of Fit". In: *Journal of the American Statistical Association* 46.253, pp. 68–78.

Matthews, C. A. and Sternlicht, D. D. (2011). "Seabed change detection in challenging environments". In: *Proceedings of the SPIE* 8017, 80170P-1 –80170P-6.

McDonald, M. A., Hildebrand, J. A., and Wiggins, S. M. (2006). "Increases in deep ocean ambient noise in the Northeast Pacific west of San Nicolas Island, California". In: *Journal of the Acoustical Society of America* 120.2, pp. 711–718.

McDonald, M. A., Hildebrand, J. A., Wiggins, S. M., and Ross, D. (2008). "A 50-year comparison of ambient ocean noise near San Clemente Island: A bathymetrically complex coastal region off southern California". In: *Journal of the Acoustical Society of America* 124.4, pp. 1985–1992.

McKenna, M. F., Ross, D., Wiggins, S. M., and Hildebrand, J. A. (2012). "Underwater radiated noise from modern commercial ships". In: *Journal of the Acoustical Society of America* 131.1, pp. 92–103.

McKenna, M. F., Soldevilla, M., Oleson, E., Wiggins, S., and Hildebrand, J. A. (2009). "Increased underwater noise levels in the Santa Barbara channel from commercial ship traffic and the potential impact on blue whales (*Balaenoptera Musculus*)". In: *Proceedings of the 7th California Islands Symposium*. Ed. by C. C. Damiani and D. K. Garcelon. Arcata, California.

Michalopoulos, D. (1979). "New Applications & Recent Research". In: *Computer* 12.3, pp. 118–119.

Michel, R. and Rignot, E. (1999). "Flow of Glaciar Moreno, Argentina, from repeat-pass Shuttle Imaging Radar images: comparison of the phase correlation method with radar interferometry". In: *Journal of Glaciology* 45.149, pp. 93–100.

Midtgaard, Ø. (2013). "Change detection using synthetic aperture sonar imagery with variable time intervals". In: *Proceedings of the 1st Underwater Acoustics Conference*. Corfu, Greece, pp. 713–720.

Midtgaard, Ø., Hansen, R. E., Sæbø, T. O., Myers, V., Dubberley, J. R., and Quidu, I. (2011). "Change detection using Synthetic Aperture Sonar: Preliminary results from the Larvik trial". In: *OCEANS 2011*. Waikoloa, HI.

Misra, P. and Enge, P. (2006). *Global positioning system: signals, measurements, and performance*. 2nd ed. Ganga-Jamuna Press.

Moore, A. J., Hand, D. P., Barton, J. S., and Jones, J. D. C. (1999). "Transient deformation measurement with electronic speckle pattern interferometry and a high-speed camera". In: *Applied Optics* 38.7, pp. 1159–1162.

Moore, E. H. (1920). "On the reciprocal of the general algebraic matrix". In: *Bulletin of the American Mathematical Society* 26.9, pp. 394–395.

Munafò, A., Sliwka, J., Ferri, G., Vermeij, A., Goldhahn, R., LePage, K., Alves, J., and Potter, J. (2014). "Enhancing AUV localization using underwater acoustic sensor networks: Results in long baseline navigation from the COLLAB13 sea trial". In: *2014 OCEANS - St. John's*.

Myers, V., Fortin, A., and Simard, P. (2009). "An automated method for change detection in areas of high clutter density using sonar imagery". In: *Proceedings of the UAM 2009 Conference*.

Myers, V., Quidu, I., Sæbø, T. O., and Hansen, R. E. (2013). "Results and analysis of coherent change detection experiments using repeat-pass synthetic aperture sonar images". In: *Proceedings of the 1st Underwater Acoustics Conference*. Corfu, Greece, pp. 591–598.

Myint, S. W., Chandra, G., Wang, L., Zhu, Z., and Gillette, S. C. (2008). "Identifying Mangrove Species and Their Surrounding Land Use and Land Cover Classes Using an Object-Oriented Approach with a Lacunarity Spatial Measure". In: *GIScience & Remote Sensing* 45.2, pp. 188–208.

Nakamura, Y., Yamaguchi, I., and Hama, Y. (2004). "Buried object detection with Synthetic Aperture Sonar". In: *Proceedings of the 2004 International Symposium on Underwater Technology*, pp. 27–31.

National Research Council (2003). *Ocean Noise and Marine Mammals*. Washington, DC: The National Academies Press.

Neilson, R. O. (1991). *Sonar Signal Processing*. Artech House.

Nelson, J. D. B. and Kingsbury, N. G. (2012). “Fractal dimension, wavelet shrinkage and anomaly detection for mine hunting”. In: *IET Signal Processing* 6.5, pp. 484–493.

Nelson, J. D. B. and Krylov, V. (2014). “Textural lacunarity for semi-supervised detection in sonar imagery”. In: *IET Radar, Sonar and Navigation* 8.6, pp. 616–621.

Nystuen, J. A. (1986). “Rainfall measurements using underwater ambient noise”. In: *The Journal of the Acoustical Society of America* 79.4, pp. 972–982.

Nystuen, J. A. and Farmer, D. M. (1987). “The influence of wind on the underwater sound generated by light rain”. In: *The Journal of the Acoustical Society of America* 82.1, pp. 270–274.

Oliver, B. M. (1963). “Sparkling spots and random diffraction”. In: *Proceedings of the IEEE* 51.1, pp. 220–221.

Oliver, C. J., McConnell, I., and Corr, D. G. (1999). “Multitemporal change detection for SAR imagery”. In: *Proceedings of the SPIE* 3869, pp. 55–66.

Oppenheim, A. V. and Lim, J. S. (1981). “The Importance of Phase in Signals”. In: *Proceedings of the IEEE* 69.5, pp. 529–541.

Orfanidis, S. J. (1989). *Optimum Signal Processing*. 2nd ed. Macmillan.

Osborn, J., Föhring, D., Dhillon, V. S., and Wilson, R. W. (2015). “Atmospheric Scintillation in Astronomical Photometry”. In: *Monthly Notices of the Royal Astronomical Society* 452.2, pp. 1707–1716.

Parker, J. A., Kenyon, R. V., and Troxel, D. E. (1983). “Comparison of Interpolating Methods for Image Resampling”. In: *IEEE Transactions on Medical Imaging* 2.1, pp. 31–39.

Pearson, K. (1895). “Notes on regression and inheritance in the case of two parents”. In: *Proceedings of the Royal Society of London* 58, pp. 240–242.

Penrose, R. (1955). “A generalized inverse for matrices”. In: *Proceedings of the Cambridge Philosophical Society* 51.3, pp. 406–413.

Perona, P. and Malik, J. (1990). “Scale-space and edge detection using anisotropic diffusion”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 12.7, pp. 629–639.

Persons, C. M., Chenault, D. B., Jones, M. W., Spradley, K. D., Gulley, M. G., and Farlow, C. A. (2002). “Automated registration of polarimetric imagery using Fourier transform techniques”. In: *Proceedings of the SPIE* 4819, pp. 107–117.

Petzing, J. N. and Tyrer, J. R. (1998). "Recent developments and applications in electronic speckle pattern interferometry". In: *The Journal of Strain Analysis for Engineering Design* 33.2, pp. 153–169.

Pilbrow, E. N. (2007). "Synthetic Aperture Sonar Micronavigation Using An Active Acoustic Beacon". PhD thesis. University of Canterbury.

Plotnick, R. E., Gardner, R. H., Hargrove, W. W., Prestegaard, K., and Perlmutter, M. (1996). "Lacunarity analysis: A general technique for the analysis of spatial patterns". In: *Physical Review E* 53.5, pp. 5461–5468.

Poeckert, R. H. (1991). *Change detection using "blink" comparison of route survey sonar imagery*. Tech. rep. DREP-TM-91-21. Victoria BC, Canada: Defence Research Establishment Pacific.

Preiss, M. and Stacy, N. J. S. (2006). *Coherent change detection: theoretical description and experimental results*. Tech. rep. DSTO-TR-1851. Australia: Defence Science and Technology Organisation.

Quazi, A. H. (1981). "An overview on the time delay estimate in active and passive systems for target localization". In: *IEEE Transactions on Acoustics, Speech, and Signal Processing* 29.3, pp. 527–533.

Quidu, I., Myers, V., Midtgaard, Ø., and Edgar Hansen, R. (2012). "Subpixel image registration for coherent change detection between two high resolution sonar passes". In: *Proceedings of the International Conference on Underwater Remote Sensing*. Brest, France, pp. 207–226.

Quinn, R. (2014). "Acoustic Remote Sensing in Maritime Archaeology". In: *The Oxford Handbook of Maritime Archaeology*. Ed. by A. Catsambis, B. Ford, and D. L. Hamilton. Oxford University Press. Chap. 3, pp. 68–89.

Rabelo, G. F., Braga Júnior, R. A., Fabbro, I., Trivi, M. R., Rabal, H. J., and Arizaga, R. (2005). "Laser speckle techniques in quality evaluation of orange fruits". In: *Revista Brasileira de Engenharia Agrícola e Ambiental* 9.4, pp. 570–575.

Radke, R. J., Andra, S., Al-Kofahi, O., and Roysam, B. (2005). "Image Change Detection Algorithms: A Systematic Survey". In: *IEEE Transactions on Image Processing* 14.3, pp. 294–307.

Ramachandran, G. N. (1943). "Fluctuations of light intensity in coronae fomed by diffraction". In: *Proceedings of the Indian Academy of Sciences, Section A* 18.4, pp. 190–200.

Lord Rayleigh (1880). "On the Resultant of a large Number of Vibrations of the same Pitch and of arbitrary Phase". In: *Philosophical Magazine Series 5* 10.60, pp. 73–78.

Lord Rayleigh (1896). *The Theory of Sound, Volume 2*. 2nd ed. Macmillan.

Lord Rayleigh (1918). "On the Light emitted from a Random Distribution of Luminous Sources". In: *Philosophical Magazine Series 6* 36.216, pp. 429–449.

Lord Rayleigh (1919). "On the Problem of Random Vibrations, and of Random Flights in one, two, or three Dimensions". In: *Philosophical Magazine Series 6* 37.220, pp. 321–346.

Rigden, J. D. and Gordon, E. I. (1962). "The granularity of scattered optical maser light". In: *Proceedings of the IRE* 50, pp. 2367–2368.

Rignot, E. J. M. and Zyl, J. J. van (1993). "Change detection techniques for ERS-1 SAR data". In: *IEEE Transactions on Geoscience and Remote Sensing* 31.4, pp. 896–906.

Roderick, W. I., Dullea, R. K., and Syck, J. M. (1984). "High resolution bottom backscatter measurements". In: *The Journal of the Acoustical Society of America* 75.S31.

Rodriguez, E. and Martin, J. M. (1992). "Theory and design of interferometric synthetic aperture radars". In: *IEEE Proceedings F — Radar and Signal Processing* 139.2, pp. 147–159.

Ross, D. (1976). *Mechanics of Underwater Noise*. New York: Pergamon.

Sæbø, T. O., Hansen, R. E., Callow, H. J., and Synnes, S. A. V. (2011). "Coregistration of synthetic aperture sonar images from repeated passes". In: *Proceedings of the 4th International Conference and Exhibition on Underwater Acoustic Measurements: Technologies & Results*. Ed. by J. S. Papadakis and L. Bjørnø. Kos, Greece, pp. 529–536.

Scheiber, R. and Moreira, A. (2000). "Coregistration of interferometric SAR images using spectral diversity". In: *IEEE Transactions on Geoscience and Remote Sensing* 38.5, pp. 2179–2191.

Scrimger, J. A., Evans, D. J., McBean, G. A., Farmer, D. M., and Kerman, B. R. (1987). "Underwater noise due to rain, hail, and snow". In: *Journal of the Acoustical Society of America* 81.1, pp. 79–86.

Seymour, M. S. and Cumming, I. G. (1994). "Maximum likelihood estimation for SAR interferometry". In: *International Geoscience and Remote Sensing Symposium (IGARSS '94)*. Vol. 4, pp. 2272–2275.

Shapiro, L. H. and Burns, J. J. (1973). "Satellite observations of sea ice movement in the Bering Strait region". In: *Climate of the Arctic: Twenty-fourth Alaska Science*

Conference, August 15 to 17, 1973. Ed. by G. Weller and S. A. Bowling. Geophysical Institute, University of Alaska.

Sharp, B. (1989). "Electronic speckle pattern interferometry (ESPI)". In: *Optics and Lasers in Engineering* 11.4. Special Issue on Holography and Speckle Metrology, pp. 241–255.

Shiavi, R. (2010). *Introduction to Applied Statistical Signal Analysis*. 3rd ed. Elsevier Science.

Singh, A. (1989). "Digital change detection techniques using remotely-sensed data". In: *International Journal of Remote Sensing* 10.6, pp. 989–1003.

Skolnik, M. I. (1980). *Introduction to radar systems*. 2nd ed. New York: McGraw-Hill.

Skolnik, M. I. (1990). *Radar Handbook*. Electronic Engineering series. McGraw-Hill.

Smirnov, N. (1939a). "On the estimation of the discrepancy between critical curves of distribution of two independent samples". In: *Bulletin Mathématique de l'Université de Moscou* 2, pp. 1–16.

Smirnov, N. (1939b). "Sur les écarts de la courbe de distribution empirique [On the deviations of the empirical distribution curve]". In: *Recueil Mathématique* 6, pp. 3–26.

Smirnov, N. (1948). "Table for estimating the goodness of fit of empirical distributions". In: *The Annals of Mathematical Statistics* 19, pp. 279–281.

Solberg, A. H. S. and Jain, A. K. (1997). "Texture Fusion and Feature Selection Applied to SAR imagery". In: *IEEE Transactions on Geoscience and Remote Sensing* 35.2, pp. 475–479.

Soumekh, M. (1994). *Fourier Array Imaging*. Prentice-Hall.

Sternlicht, D. D., Fernandez, J. E., Prater, J. L., Weaver, J. N., Isaacs, J. C., Montgomery, T. C., Loeffler, C. M., and Purcell, M. (2016). "Advanced Sonar Technologies for High Clearance Rate Mine Countermeasures". In: *OCEANS 2016 MTS/IEEE Monterey*.

Sternlicht, D. D., G-Michael, T., and Matthews, C. A. (2012). "Advances in seabed change detection for port and coastal security". In: *Proceedings of the 3rd International Conference on Waterside Security*.

Sternlicht, D. D., Harbaugh, J. K., and Nelson, M. A. (2009). "Experiments in coherent change detection for synthetic aperture sonar". In: *OCEANS 2009, MTS/IEEE Biloxi - Marine Technology for Our Future: Global and Local Challenges*. Biloxi, MS.

Stolt, R. H. (1978). “Migration by Fourier transform”. In: *Geophysics* 43.1, pp. 23–48.

Synnes, S. A. V., Callow, H. J., Hansen, R. E., and Sæbø, T. O. (2010). “Multipass coherent processing on synthetic aperture sonar data”. In: *Proceedings of European Conference on Underwater Acoustics (ECUA) 2010*. Istanbul, Turkey.

Théau, J. (2008). “Change Detection”. In: *Encyclopedia of GIS*. Ed. by S. Shekhar and H. Xiong. Springer, pp. 77–84.

Therrien, C. W. (1999). “Overview of Statistical Signal Processing”. In: *Digital Signal Processing Handbook on CD-ROM*. Ed. by V. Madisetti and D. Williams. CRC Press. Chap. 12.

Tinkle, A. D. and Chang, E. (2001). “Synthetic aperture sonar point response for buried objects”. In: *MTS/IEEE Oceans 2001: An Ocean Odyssey*. Vol. 1, pp. 225–233.

Tomiyasu, K. (1978). “Tutorial review of synthetic-aperture radar (SAR) with applications to imaging of the ocean surface”. In: *Proceedings of the IEEE* 66.5, pp. 563–583.

Touzi, R. and Lopes, A. (1996). “Statistics of the Stokes parameters and of the complex coherence parameters in one-look and multilook speckle fields”. In: *IEEE Transactions on Geoscience and Remote Sensing* 34.2, pp. 519–531.

Touzi, R., Lopes, A., Bruniquel, J., and Vachon, P. W. (1999). “Coherence estimation for SAR imagery”. In: *IEEE Transactions on Geoscience and Remote Sensing* 37.1, pp. 135–149.

Tucker, J. D. and Azimi-Sadjadi, M. R. (2011). “Coherence-Based Underwater Target Detection From Multiple Disparate Sonar Platforms”. In: *IEEE Journal of Oceanic Engineering* 36.1, pp. 38–52.

University of California, Division of War Research (1946). *Underwater noise caused by snapping shrimp*. Tech. rep. Contract NObs-2074 (formerly OEMsr-30). Navy Department, Bureau of Ships.

Urick, R. J. (1975). *Principles of underwater sound*. 2nd ed. New York: McGraw-Hill.

Venot, A., Golmard, J. L., Lebruchec, J. F., Pronzato, L., Walter, E., Frij, G., and Roucayrol, J. C. (1984). “Digital methods for change detection in medical images”. In: *Information processing in medical imaging*. Springer, pp. 1–16.

Viola, P. and Jones, M. (2004). “Robust real-time object detection”. In: *International Journal of Computer Vision* 57.2, pp. 137–154.

Vossen, R. van, B. A. J. Quesson, A. L. D. Beckers, M. Zampolli, M. E. G. D. Colin, and A. J. Hunter (2012). “Low-frequency synthetic aperture sonar system for the

detection of objects buried in mud". In: *Proc. Undersea Defence Technology (UDT) Europe*. Alicant, Spain.

Wang, V. and Hayes, M. P. (2014). "Image registration of simulated synthetic aperture sonar images using SIFT". In: *Proceedings of the 29th International Conference on Image and Vision Computing New Zealand*. IVCNZ '14. Hamilton, New Zealand, pp. 31–36.

Wang, V. and Hayes, M. P. (2016a). "Analysis of feature matching performance on correlated speckle image pairs". In: *OCEANS 2016 MTS/IEEE*. Monterey.

Wang, V. and Hayes, M. P. (2016b). "Modelling of feature matching performance on correlated speckle images". In: *2016 International Conference on Image and Vision Computing New Zealand (IVCNZ)*. Palmerston North, New Zealand.

Wang, V. and Hayes, M. P. (2017). "Synthetic Aperture Sonar Track Registration Using SIFT Image Correspondences". In: *IEEE Journal of Oceanic Engineering*. In press.

Williams, D. P. (2015). "Fast Unsupervised Seafloor Characterization in Sonar Imagery Using Lacunarity". In: *IEEE Transactions on Geoscience and Remote Sensing* 53.11, pp. 6022–6034.

Willumsen, A. B., Hallingstad, O., and Jalving, B. (2006). "Integration of Range, Bearing and Doppler Measurements from Transponders into Underwater Vehicle Navigation Systems". In: *OCEANS 2006*.

Wilson Jr, O. B., Wolf, S. N., and Ingenito, F. (1985). "Measurements of acoustic ambient noise in shallow water due to breaking surf". In: *The Journal of the Acoustical Society of America* 78.1, pp. 190–195.

Wolfschmidt, G. (1998). "Stereo and blink comparators". In: *Instruments of Science: An Historical Encyclopedia*. Ed. by R. Bud and D. J. Warner. Taylor and Francis, pp. 126–127.

Zebker, H. A. and Villasenor, J. (1992). "Decorrelation in interferometric radar echoes". In: *IEEE Transactions on Geoscience and Remote Sensing* 30.5, pp. 950–959.